Applications of Wavelet Transforms to Analysing Medical Signals

Katherine Elizabeth Macey

A thesis presented for the degree of
Doctor of Philosophy
in
Electrical and Electronic Engineering
at the
University of Canterbury,
Christchurch, New Zealand.

December, 2000

ABSTRACT

In this thesis, wavelets are applied to the analysis of two types of medical signals, namely infant breathing signals and ultrasound images. One-dimensional wavelets are used to quantify amplitude modulation of infant breathing that occurs during quiet sleep. Two-dimensional wavelets are used to develop enhancement techniques in the wavelet domain, tailored to ultrasound images.

The development of wavelets is described from the backgrounds of mathematics, signal processing and sub-band coding. One-dimensional wavelets are defined for both continuous and discrete cases. Two-dimensional wavelets are developed for both separable and non-separable classes. The construction of both one- and two-dimensional wavelets is described, and examples of wavelets that are used in subsequent analyses are presented.

The analysis of breathing signals provides information for understanding the physiology of breathing. The wavelet domain is shown to isolate frequency characteristics of breathing signals, as well as indicating the temporal position of those characteristics in the signal. Some of the characteristics that were measured were not distinguished in the original signals. The extent of constant frequency components due to amplitude modulation of the principal breathing rate is quantified. Breathing in infants is of particular interest, as it may provide insight into the cause of Sudden Infant Death Syndrome (SIDS), or cot death. Studies of infants who later succumbed to SIDS, or were at high risk for SIDS, and infants at low risk for SIDS were carried out. The infants who later succumbed to SIDS and those at high risk for SIDS showed different characteristics in the wavelet domain compared to infants at low risk for SIDS. As well as implying that there may be a difference in the physiologies of infants at high and low risk for SIDS, this result confirms that wavelets can be used to analyse breathing signals and produce meaningful results.

Ultrasound images typically contain artefacts and low contrast between features of interest and often exhibit a noisy background. The aim of any enhancement procedure is to reduce the contributions from noise and artefacts and increase the contrast between the features of interest and the background. A method to objectively measure improvements in image quality is described, although this attempt achieved mixed results. Separable and non-separable two-dimensional wavelets are used with

enhancement functions to improve the image quality. Enhancement schemes, including noise reduction and contrast enhancement, are developed. A spatially varying contrast enhancement function is also developed. Noise reduction, in combination with the spatially varying contrast enhancement function, produces an image that reduces the existing artefacts in the image and increases the contrast for the features of interest.

Wavelets are shown to be a useful tool for the analysis of infant breathing signals and for improvement of image quality in ultrasound images.

PREFACE

This preface describes the history of this thesis, and gives an executive summary of the contents.

The research presented in this thesis is concerned with the development and application of engineering techniques to medical research. There are two areas of medical research that have been considered: the investigation of breathing in relation to Sudden Infant Death Syndrome (SIDS), or Cot Death, and the enhancement of ultrasound images.

I was introduced to the world of signal processing by the late Professor R.H.T. Bates. Professor Bates was a prolific and much respected researcher and lecturer at the University of Canterbury until his death in 1990. Signal processing caught my imagination during a course I took of his in my final undergraduate year. It was not until mid-way through 1992 that I followed that path.

Dr. Kathy Garden was my first supervisor, and, at the time, was setting up a project, in conjunction with Lincoln University and Christchurch Hospital, imaging sheep using a Computed Tomography (CT) scanner. It was through her that I met Dr. Richard Fright, who was then working in the Department of Medical Physics and Bioengineering at Christchurch Hospital. When Dr. Garden moved, it was Dr. Fright who continued to provide the support and direction for my thesis. Dr. Pat Bodger joined Dr. Fright in supervising me, providing me with administrative and general support. The focus of my research moved from CT to ultrasound images. I became interested in wavelets as a tool to analyse ultrasound images.

In 1996, I developed a strong friendship with another Ph.D. student in the Electrical and Electronic Engineering Department, Dr. Paul Macey. Through this contact I saw some of the breathing data that had been collected by the BabyLog cot death research group headed by Dr. Rodney Ford, a community paediatrician. During 1985, Dr. Ford recognised the need for the application of engineering techniques to help study infant physiology and clinical aspects of SIDS. Thus, Dr. Ford, the late Professor R.H.T. Bates, Dr. Fright, and Dr. Garden began a collaborative research effort between the Christchurch Hospital, and the Department of Electrical and Electronic Engineering at the University of Canterbury. To assist in the study of infant physiology, a system now known as *BabyLog* was developed to collect and store infant physiological signals.

Since then, various people have been involved in collecting and analysing these signals, and the group is now known as the BabyLog cot death research group. My research into the characteristics of breathing signals evolved from this group project.

In September 1998, a break in the research preceeded a move to Massey University in July 1999 to work with Dr. Wyatt Page for a month, before moving to Los Angeles and working with Dr. Ronald Harper at the University of California at Los Angeles (UCLA) until July 2000. Drs. Page, Harper and Bodger continued to supervise my work in the last year.

The following paragraphs outline the structure of this thesis and identify original areas of my research. Parts of this research have been conducted in collaboration with others. Original and collaborative efforts have been made clear where appropriate.

Chapter 1 provides the motivation and introduction to the use of wavelets for analysing signals. The problem of SIDS is introduced and briefly reviewed. The relevance of wavelets research to SIDS is highlighted. Ultrasound imaging is also reviewed. The role of wavelets for analysis is presented. No original work is presented in this chapter.

Chapter 2 introduces the field of wavelet transforms in more detail. The development of the field is reviewed, providing background information for the remainder of the thesis. Again, no original work is presented in this chapter.

Chapter 3 describes the development of one-dimensional wavelets, and their construction. No original work is presented in this chapter.

Chapter 4 describes the characteristics of breathing signals leading to a novel method, the Wavelet Performance Measure, for the selection of a wavelet to analyse amplitude modulation of infant breathing signals. This method for selection of a wavelet is entirely new and my own work. The direction of research was decided in collaboration with Dr. Page.

The results of applying wavelets to infant breathing signals are presented in Chapter 5. Two novel measures of amplitude modulation in the wavelet domain, rhythmicity and impact, are developed. This chapter consists of entirely new efforts and my own work. The direction of research was decided in collaboration with Drs. Macey, Ford, Fright, Harper and Page.

Chapter 6 describes the construction of two-dimensional separable and non-separable wavelets. Results of applying wavelets in two dimensions are presented with descriptions of features and characteristics in the wavelet domain. No original work is presented in this chapter.

Image quality is addressed in Chapter 7 and a novel method for assessing the image quality of ultrasound images is introduced. The development of the novel image quality equation is my own work. The direction of research was decided in collaboration with Dr. Page.

Chapter 8 examines image enhancement techniques for ultrasound images for use in the wavelet domain. The enhancement techniques are broadly classified as noise reduction and contrast enhancement. Two methods for noise reduction are developed: an extension of a method presented by Roy *et al.* [1999] to two-dimensions, Two-Dimensional Derivative Denoising, and a novel method, Wavelet Median Denoising. Two functions developed for contrast enhancement in the wavelet domain are based on existing functions presented by Laine *et al.* [1995] and Chang and Vetterli [1997], which were adapted and extended for application to ultrasound images. There was collaboration with Drs. Fright, Page and Mr. Nigel Anderson on the interpretation and assessment of results and direction of research. Results of applying the image enhancement techniques to ultrasound images are presented. These results are my own work, with direction from Drs. Fright and Page.

Chapter 9 presents conclusions regarding the original contributions in this thesis. The applicability of wavelet transforms as an analysis tool for infant breathing signals is discussed. The conclusions regarding the use of wavelet transforms in conjunction with enhancement techniques for ultrasound images are presented. Finally, suggestions for further work are given.

The following are the publications and presentations prepared during the course of my Ph.D research:

Surman, K.[1], Carr, J., Columbi, Y., Fright, R., and Garden, K. (1993), "Development of a System for Three-dimensional Graphics from Ultrasound," In *Proceedings of the First New Zealand Conference on Image and Vision Computing*, Auckland, New Zealand, 16–18 August, 1993, pp. 367-373.

Macey, K.E., L. Kuo, A. Kim, P.L. Yu, M.A. Woo, M.M. Saeed, D. Gozal, and R.M. Harper, Breathing Patterns Following Exposure to Carbon Dioxide in Congenital Central Hypoventilation Syndrome, *Sleep*, 14th Annual Associated Professional Sleep Societies Meeting, 17–22 June, 2000, Las Vegas, Nevada. Vol 23, No. 2, 2000, pp A16–17.

Macey, K.E., R.M. Harper, W.H. Page, and P.M. Macey, Quantifying Amplitude Modulation of Breathing in Infants Using the Wavelet Transform, In *CD-ROM Proceedings of the World Congress on Medical Physics and Biomedical Engineering*, 23–28 July, 2000, paper 5892-11144, 4 pages, 2000.

Macey, K.E., W.H. Page, R.M. Harper, P.M. Macey, and R.P.K. Ford. Calculating Rhythmicity of Infant Breathing Using Wavelets, In *Proceedings on SPIE Conference on Wavelet Applications in Signal and Image Processing VIII*, 30 July–4 August, 2000. Vol. 4119, December 2000.

Macey, K.E. and W. Page, A Method for Selecting a Wavelet to Analyse Infant Breathing Signals. Submitted to *Applied Signal Processing.*

[1]Maiden name

In Preparation:

Macey, K.E., P.M. Macey, W.H. Page, and Harper, R.M., "Time of Night effects on Amplitude Modulation in Sudden Infant Death Syndrome and Control Infants," to be submitted to *Pediatric Research.*

Macey, K.E. and W.H. Page, "Image Enhancement of Ultrasound Images using Wavelet Transforms," to be submitted to *Ultrasound in Medicine and Biology.*

ACKNOWLEDGMENTS

There are many people who have contributed to this thesis in many ways, both professionally and personally.

Firstly, I thank my supervisors. I thank Drs. Wyatt Page and Ronald Harper for providing me with encouragement, commitment to my thesis, and valuable guidance in the last year of my thesis. I thank Dr. Richard Fright for his long support and guidance of my research and thesis. I also thank Dr. Kathy Garden for her encouragement to pursue post-graduate research, and her guidance in the initial stages of my research. I thank Dr. Pat Bodger for his prompt review of my work and his holistic encouragement and belief in me, my lifestyle and my research.

I also thank Dr. Ronald Harper for giving me the space and equipment needed for the completion of my thesis at UCLA.

I thank those people who have provided me with additional valuable information regarding medical matters. I thank Dr. Rodney Ford for his insights into infant physiology and Nigel Anderson for his insights into diagnosing ultrasound images. I admire both of them for their enthusiastic approach to research.

I thank the late Richard Cox, Dave van Leeuwen, Pieter Kikstra, Mike Shurety, and Florin Predan for their unfailing ability to happily answer all my questions and fix all my problems regarding the computer network.

I thank Carl Taswell for the free "wavebox" software, on which I based my two-dimensional wavelet transform programmes.

I thank Telecom New Zealand, Ltd. for providing me with financial support for the first three and a half years of my research.

On a personal level, I thank the staff and students of the Electrical and Electronic Engineering Department of the University of Canterbury who are my colleagues and have become my friends and mentors.

I thank Joe DiMaggio, Jane Murphy and the Christchurch Landmark Education team, and the Transformations team without whom I would have struggled to the very end. It is because of them and my interactions with them that I continued my research and enjoyed it.

I also thank my canoe polo team mates, local and national, and my whitewater paddling buddies, with whom I've travelled to places on rivers and in the outdoors and the world that hold the utmost inspiration for being alive.

I thank Sri Chinmoy and members of the Sri Chinmoy meditation group with whom I have gained immeasurable spiritual guidance and relaxation in the last three and a half years.

Lastly, I thank those who are most important to me:

Paul Macey for both his love, companionship and support as my husband, and also his direction and feedback of my research and my writing.

My parents, Liz and Chris, and the rest of my family, for providing me with unconditional love, encouragement, and support.

CONTENTS

GLOSSARY

PARAMETERS

λ	Noise reduction threshold
$\phi(t)$	Scaling function
$\Phi(x,y)$	Scaling function in 2-D
$\psi(t)$	Wavelet function
ψ^N	Daubechies orthogonal wavelets, order N
$\Psi(f)$	Fourier Transform of wavelet function
$\Psi(x,y)$	Wavelet function in 2-D
2^j	Resolution at scale, j
a	Scale in the continuous wavelet transform
a_j	Approximation for discrete 1-D or 2-D non-separable wavelets, at j^{th} scale
$a^j_{0,0}$	Approximation for discrete 2-D separable wavelets, at j^{th} scale
A_{2^j}	Orthogonal projection operator of vector space V_{2^j}
C^n	Space of functions having n continuous derivatives
d_j	Detail for discrete 1-D or 2-D non-separable wavelets, at j^{th} scale
$d^j_{k,l}$	Detail for discrete 2-D separable wavelets, at j^{th} scale and $k,l = 0,1$
$Det(\mathbf{D})$	Determinant of the matrix $\mathbf{D}$
$\mathbf{D}$	Dilation matrix in 2-D
$E(t)$	Enhancement function
f	Frequency parameter
$g[n]$	Discrete time domain synthesis filter
$g_0[n]$	Discrete time domain low-pass synthesis filter
$g_1[n]$	Discrete time domain high-pass synthesis filter
$G_p(z)$	Polyphase synthesis filter in the z-domain
$G(z)$	Synthesis filter in the z-domain
$h[n]$	Discrete time domain analysis filter
$h_0[n]$	Discrete time domain low-pass analysis filter
$h_1[n]$	Discrete time domain high-pass analysis filter
$H_p(z)$	Polyphase analysis filter in the z-domain
$H(z)$	Analysis filter in the z-domain
$\mathcal{H}^p(\mathcal{R})$	Real Hardy spaces

PARAMETERS (CONTINUED)

i	Square root of -1
I	Indentity matrix
$\mathcal{I}^2(\mathcal{Z})$	vector space of all square-summable sequences
j	j^{th} scale in the discrete wavelet transform, $j \in \mathcal{Z}$
$\mathbf{k}$	Gain used in spatially varying contrast enhancement
$\mathbf{K}$	Gain used in contrast enhancement
$L^2(\mathcal{R}^n)$	Space of measurable, square integrable n-dimensional functions
O_{2j}	Orthogonal complement to V_{2j}
$\mathbf{O}_{2j}$	Orthogonal complement to $\mathbf{V}_{2j}$
$\mathcal{R}^n$	Euclidean space of n-dimensions
s	Statistical variance
t	Signal parameter (temporal or spatial) in 1-D
$\mathbf{t}$	Signal parameter (spatial) in 2-D
T	Contrast Enhancement Threshold
V_{2j}	A vector space that is the set of all approximations of a signal in $L^2(\mathcal{R}^1)$
$\mathbf{V}_{2j}$	A vector space that is the set of all approximations of a signal in $L^2(\mathcal{R}^2)$
$x[n]$	A discrete signal in the temporal or spatial domain
$x(t)$	A continuous signal in the temporal or spatial domain
$x'(t)$	Noise reduced signal
$\hat{x}(t)$	Output signal from a filter bank in the time domain
$X(f)$	Fourier Transform of a signal, in the frequency domain
$X(z)$	z-Transform of a signal, in the z-domain
$\mathcal{Z}$	Set of integers in 1-D
$\mathcal{Z}^n$	Set of integers in n-dimensions

ABBREVIATIONS

ALTE	Apparent Life Threatening Event
CWT	Continuous Wavelet Transform
FIR	Finite Impule Response (filters)
Hz	Hertz
MHz	Mega Hertz
QMF	Quadrature Mirror Filter
REM	Rapid Eye Movement (sleep)
SIDS	Sudden Infant Death Syndrome
STFT	Short Time Fourier Transform

MATHEMATICAL SYMBOLS

∂	Partial derivative
ϵ	Is an element of
$\forall$	For all
$\exists$	There exists
$\sum$	The sum of
$\prod$	The product of
$b \Leftrightarrow c$	b implies c and c implies b
$\mid \quad \mid$	Absolute value of
$\| \quad \|$	Magnitude of
$< \quad >$	Inner product of
$x^*(t)$	Complex conjugate of $x(t)$
$(X)^T$	Transpose of matrix X
∞	Infinity
π	pi
$\lim_{j \to \infty}$	The limit as j tends to infinity
$\odot$	Convolution
$A \oplus B$	Append the vector space, A, to the vector space, B
$\otimes$	Tensor product
$\subset$	Subset of
$\bigcup$	Union of
$\bigcap$	Intersection of
$\perp$	Orthogonal to

Chapter 1

INTRODUCTION

The aim of this thesis is to use wavelets as a method for analysing one- and two-dimensional medical signals to provide information from the signals that is not otherwise easily extracted. Specifically, the signals of interest are, in one-dimension, infant breathing recorded during sleep and, in two-dimensions, ultrasound images.

This chapter introduces wavelets and gives an overview of the theory with some examples of one-dimensional wavelets, preparing the reader for the following two chapters on the history of wavelets and their construction. Also in this chapter, an introduction to infant breathing is presented. The motivation for studying this type of physiological signal, the Sudden Infant Death Syndrome (SIDS), is discussed. The third section in this chapter discusses ultrasound images and the particular features of ultrasound images that are of interest to clinicians. A summary is given in the final section.

1.1 INTRODUCTION TO WAVELETS

Wavelet analysis is a recent signal processing technique. Wavelets as a field of study started to appear in the literature in the 1980's [Meyer 1993]. Theoretical mathematics developed by Grossman and Morlet [1984], Stromberg and Daubechies [1992] was married with the theory behind quadrature mirror filters (QMF's) to provide practical applications of wavelets [Meyer 1993]. The first algorithm for signal processing using wavelets appeared in 1989 [Mallat 1989]. A more detailed history of wavelets is presented in Chapter 2. The remainder of this section gives some examples of how wavelets have been used and introduces the key concepts of wavelet theory.

Wavelet Transforms have been used to code and analyse signals. In the field of sub-band coding, wavelets are used to transmit and store signals efficiently [Vetterli and Kovačević 1995]. The FBI use wavelets to store their database of fingerprints [Bradley and Brislawn 1994a, Bradley *et al.* 1994b]. Wavelets have been used in the compression of electrocardiograms [Anant *et al.* 1995a], in the coding of HDTV [Vetterli *et al.* 1990], and the compression of images [Lewis and Knowles 1992]. Wavelet transforms have also been used to analyse transient information in signals such as electroencephalo-

grams (EEG's), [Sun *et al.* 1993, Akay and Szeto 1995, Senhadji *et al.* 1995, Sun and Sclabassi 1998, Herrera *et al.* 1999], electrocardiograms (ECG's) [Anant *et al.* 1994], transistor turn-on characteristics [Lee and Yamamoto 1994], and seismograms [Anant and Dowla 1995b]. Other applications include the analysis of music and speech [Grossmann *et al.* 1987], fractal analysis, and the analysis of astronomical data [Meyer 1993]. Wavelets have been used in predicting time series [Tsui *et al.* 1995, Tsui *et al.* 1997] and even in the forecasting of financial trends [Aussem *et al.* 1998]. In two dimensions, wavelets have been used to analyse and de-noise various medical images [Healy, Jr. and Weaver 1992, Healy, Jr. *et al.* 1995, Laine *et al.* 1995, Richardson, Jr. 1995], Synthetic Aperture Radar (SAR), and other images [Böröczky *et al.* 1995, Hilton and Ogden 1997]. Particular applications in three dimensions include the analysis of medical data [Muraki 1993] and coding of video sequences [Kovačević and Vetterli 1993]. This thesis focuses on the analysis of medical signals in one- and two-dimensions.

Wavelet Transforms, as an analysis tool, are related to Fourier Transforms. The Fourier Transform is a method of analysing a temporal or spatial signal for its frequency components [Bloomfield 1976]. The Fourier Transform takes a signal from either the time or space domain and transforms it into the frequency domain. Location information available in the spatial or temporal domains is no longer apparent in the frequency domain. The Fourier Transform is performed on a time span, T, of measured or known data. All signals are regarded as stationary with a fundamental repeating frequency of $\frac{1}{T}$, i.e., self-repeating outside the time span, T. However, many signals are not stationary, as they contain transient components. The Fourier Transform does not distinguish between transient and stationary signals. Therefore, transient signal characteristics, which may only last for a small portion of the signal, are attributed to the entire signal.

To characterise features by both location and frequency, the Fourier Transform was modified by taking fixed length windows of the signal and performing the Fourier Transform for each window [Gabor 1946]. This technique, termed the Short Time Fourier Transform (STFT), assumes that the windowed portions of the signal are stationary. The Fourier Transform is performed for each window, giving the frequency components in that window. Temporal location is given by the position of the window in the original signal and frequency information is given by the Fourier Transform of the windowed signal. The STFT has improved localisation compared to the Fourier Transform, but limitations in the resolution of both the temporal and frequency components of the transform occur due to the fixed window size. This is explained in more detail in Section 2.2.

The developers of wavelet theory created a technique using different window lengths at different resolutions to minimise the limitations of resolution in both the temporal and the frequency dimensions of the transform [Meyer 1993]. A similar technique may be applied to signals originating in the spatial domain. The windows are described

at a particular resolution by the parameter of *scale*. The resulting transform has the number of dimensions of the original signal plus the added dimension of scale, i.e. a two-dimensional signal results in a three-dimensional transform. Wavelet analysis does not measure frequency directly, but the parameter of scale is related to frequency.

Wavelet transforms are achieved, at each resolution, by using a scaled version of the wavelets being used at other resolutions. Therefore, there are two functions associated with the wavelet transform—the wavelet function itself and a scaling function. These two functions are closely related and are described in Section 2.2.3.

When a wavelet transform is implemented, the wavelet function is equivalent to a high pass filter and the scaling function is equivalent to a low pass filter. The high and low pass filters form a Quadrature Mirror Filter pair [Mallat 1989].

The wavelet transform is implemented by the process depicted in Figure 1.1. The signal, x, is filtered by the low pass filter, h_0, and by the high pass filter, h_1, leading to two separate transformed signals in the wavelet domain. Each of these transformed signals can be sub-sampled by a factor of two according to Nyquist's theorem as they have half the bandwidth of the original signal. This is also called down-sampling. The down-sampled, high pass transformed signal contains the high frequency components of the signal and is termed the detail of the signal, d_j, at scale j. The down-sampled, low pass transformed signal contains the low frequency components of the signal and is termed the approximation to the signal, a_j, at scale j. The process of filtering the signal by the low pass filter, h_0, and the high pass filter, h_1, and then down-sampling by two gives the wavelet transform at one scale. To obtain the remaining required scales, the process is repeated using the approximation signal as an input until the final desired scale is reached.

Wavelet transforms can be implemented in either continuous form or discrete form. An example of a one-dimensional signal and its wavelet transform are shown in Figure 1.2(a). The signal consists of 30 seconds of breathing from a sleeping infant. The

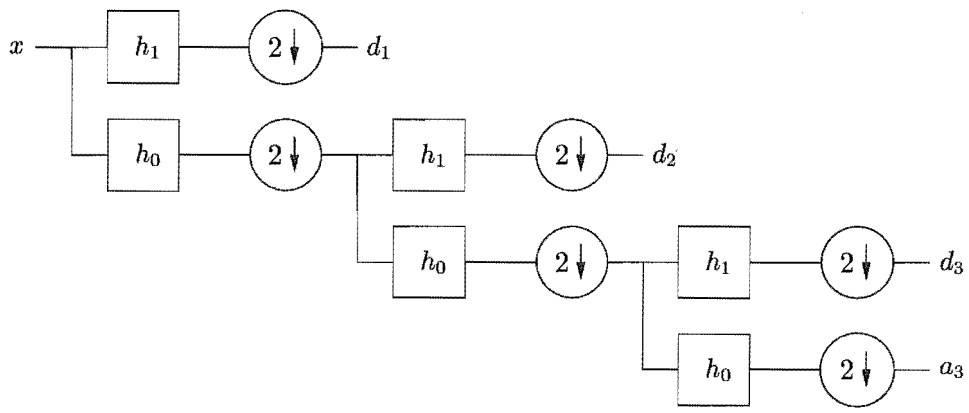
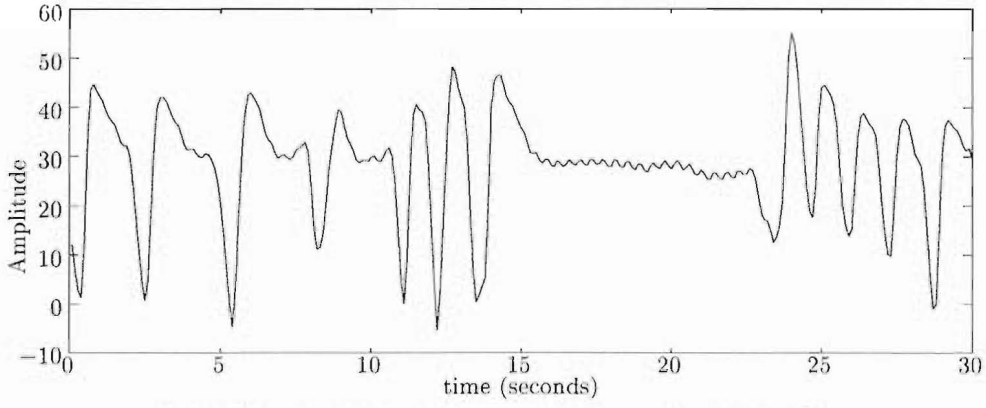


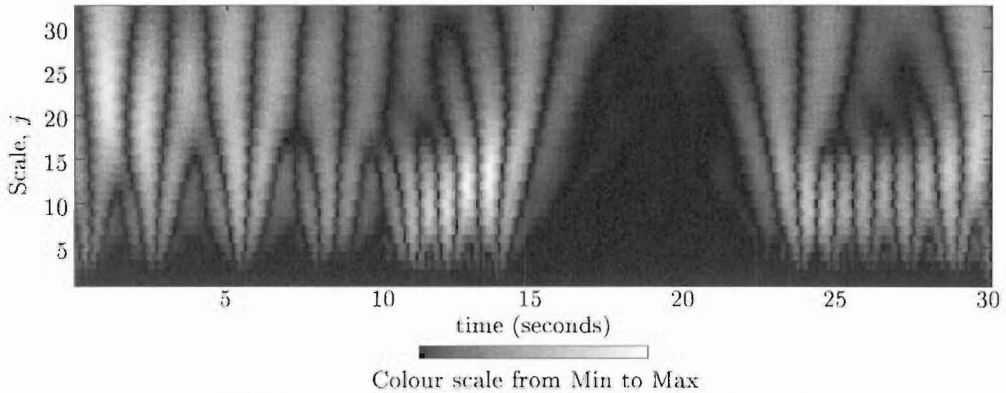
Figure 1.1 Block diagram of filter bank used to implement the wavelet transform to three scales.

features of interest are the initial regular breathing (0 10 seconds) followed by a short period of rapid breathing (10 15 seconds), followed by a period of paused breath (15–23 seconds), and finally rapid regular breathing (23 30 seconds). The continuous wavelet transform results in a two-dimensional signal over time and scale, Figure 1.2(b). The different features of the signal are highlighted in the wavelet domain at the same location as they occur in the time domain. Fan-shaped features occur during the breathing, with the rapid breathing (10 15 seconds and 23 30 seconds) having a higher intensity towards the bottom of the continuous wavelet transform than the slower breathing (0 10 seconds). The slower breathing has a higher intensity towards the top of the continuous wavelet transform. The pause in breathing (15 23 seconds) is apparent as a dark area in the continuous wavelet transform. The high frequency components of the signal are highlighted at low scales (bottom of the plot) while the low frequency components are isolated at high scales (top of the plot). Finally, the regular breathing frequency components (0- 10 seconds at the top of the plot and 11 14 seconds and 24- 29 seconds at the bottom of the plot) are seen as a series of vertical, high contrast stripes, alternating dark and light.

The same breathing signal is used to illustrate the discrete wavelet transform, as



(a) Original signal of breathing recorded from a sleeping infant.



(b) Continuous wavelet transform of breathing signal.

Figure 1.2 A signal and its Continuous Wavelet Transform.

shown in Figure 1.3. The Wavelet Transform in its discrete form in Figure 1.3(b) has individual scales shown separately. Again, the high frequency information is highlighted at low scales and the low frequency information is isolated at high scales. The rapid breathing (10-15 seconds and 23-30 seconds) is most clearly distinguished by scales, $j = 1$ and $j = 2$. The slower breathing is most clearly distinguished at scales, $j = 2$ and $j = 3$. The low frequency information regarding the presence of breathing just prior to the pause in breathing and just after the pause in breathing is most apparent in scale $j = 4$. The pause in breathing is apparent in scales $j = 2$ to $j = 4$.

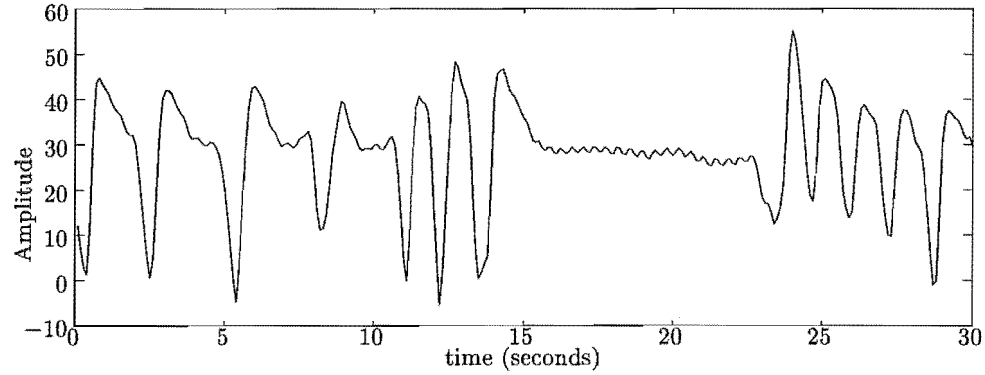
This thesis examines the use of wavelet transforms in the analysis of infant breathing signals and ultrasound images. Two-dimensional wavelet transforms are introduced in Chapter 6. The motivation for analysing sleeping infant's breathing signals and two-dimensional medical images is introduced in the next two sections.

1.2 THE ANALYSIS OF INFANT BREATHING SIGNALS

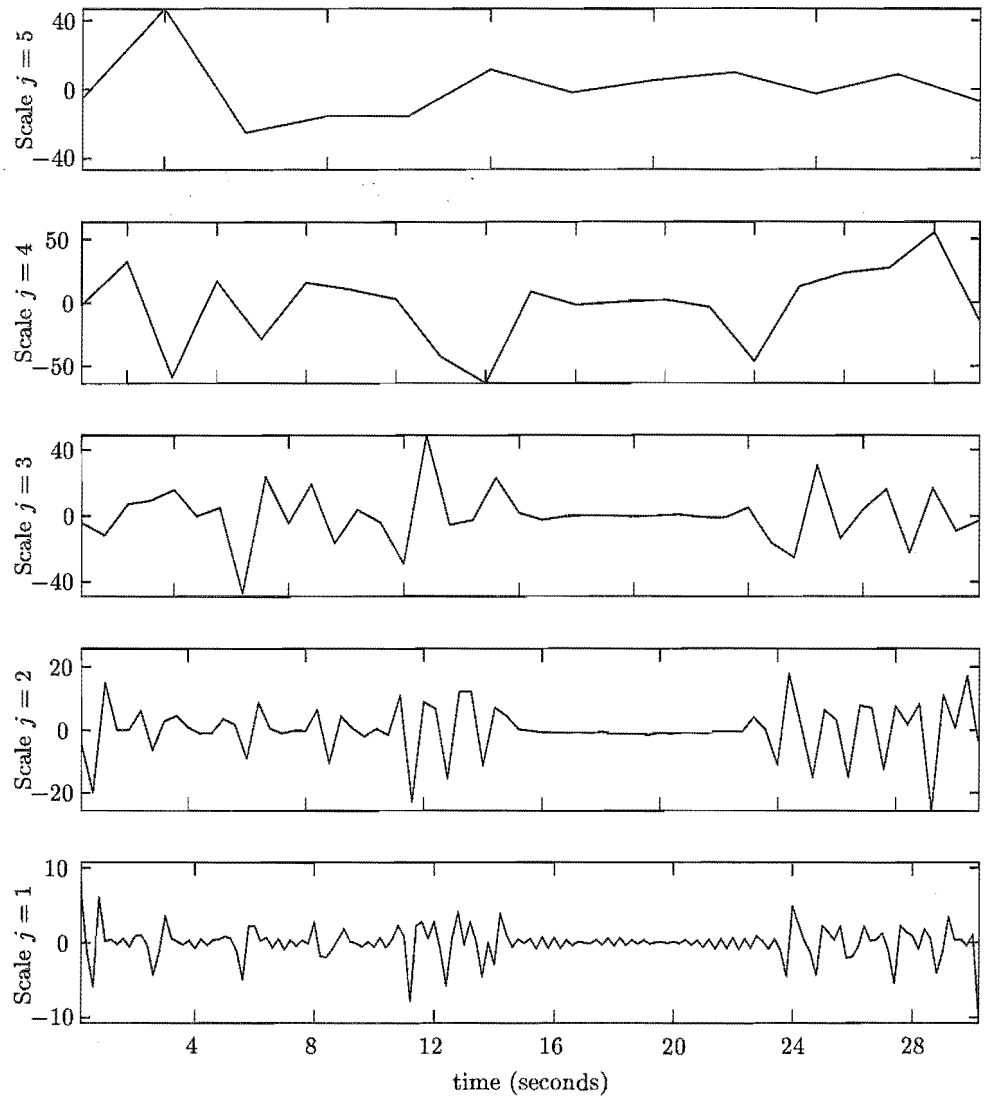
The aim of studying breathing, along with other physiological parameters, is to understand human physiology and therefore be able to diagnose, predict and prevent ailments. This section outlines the motivation for studying breathing, and in particular, the breathing of sleeping infants.

One group of patients that are seen at hospitals is infants who have experienced Near Miss Cot Deaths, or, Apparent Life Threatening Events (ALTE's) [Ariagno *et al.* 1983, Brooks 1992]. ALTE's occur during sleep and these infants have been revived after being found clammy-skinned, blue and not breathing. Infants who have suffered an ALTE are at greater risk of dying from Cot Death [Oren *et al.* 1986, Ford *et al.* 1992]. Cot Death is medically known as Sudden Infant Death Syndrome (SIDS) [Ford 1986, Hunt and Brouillette 1987, Nelson *et al.* 1989, Gibson 1996b]. SIDS occurs in infants between the ages of 4 weeks and 1 year [Gibson 1996b]. There is no apparent cause of death; the infant simply dies during sleep. While there is no known cause of SIDS, international research aims to understand the physiology of infants, and then investigate specific mechanisms that could lead to SIDS [Stein *et al.* 1979, Hunt *et al.* 1985, Lee *et al.* 1987]. A second aim of research is to discover measurable characteristics that predict SIDS, thereby allowing preventative interventions.

A research team at Christchurch Hospital is pursuing these aims. To understand the phenomenon of SIDS, infants who come into Christchurch Hospital have several physiological features measured during sleep. The signals produced from these measurements are collectively called polysomnographic recordings [Stein and Shannon 1975, Dove *et al.* 1990]. Polysomnographic recordings may include several types of breathing signals, several temperature readings, cardiac signals, blood oxygen saturation, and other variables such as body position. When assessing medical conditions other than ALTE, different variables may be measured, such as oesophageal pH for



(a) Original signal of breathing recorded from a sleeping infant.



(b) Discrete wavelet transform of breathing signal.

Figure 1.3 A signal and its Discrete Wavelet Transform.

gastro-oesophageal reflux [Dove *et al.* 1990]. However, the most commonly measured signal for all infants regardless of the condition is breathing [Macey 1998]. It is a common hypothesis throughout the world that SIDS is related to the failure of the cardio–respiratory system [Guilleminault *et al.* 1981, Hodgman *et al.* 1982, Gordon *et al.* 1984, Dunne *et al.* 1986, Kelly *et al.* 1986, Oren *et al.* 1986, Southall 1988, Milne and Ruggins 1989, Kahn *et al.* 1992, Gibson 1996b, Katz Solomon and Milerad 1996, Scheffer *et al.* 1996]. Therefore, it is logical to record and analyse breathing signals.

One type of breathing signal used in this thesis is the *Graseby* breathing signal. The Graseby breathing signal is so called because the Phillips Graseby MR10 respiration monitor is used to collect the signal [Graseby Medical Ltd. 1988]. Similar types of instrument are used internationally [Hewertson *et al.* 1944, Gordon *et al.* 1986, Schechtman *et al.* 1988]. The Graseby instrument is readily available in New Zealand and is widely used in hospitals and homes [Dove *et al.* 1990, Ford *et al.* 1992, Macey *et al.* 1995]. Due to its widespread use, a large database of breathing signals is available [Macey 1998]. It is for these reasons that Graseby breathing signals were chosen for one of the studies in this thesis.

A schematic of the Graseby instrument is shown in Figure 1.4, and consists of a capsule connected via a hollow plastic tube to a small case. The capsule is taped onto the abdominal wall of the infant. The pressure change in the capsule due to the abdominal wall movements during breathing causes the pressure to change in the tube, giving an indirect measure of the infant's breathing in normal situations. The pressure change in the tube causes the moveable plate in the capacitor to move, generating an electrical signal. This signal is recorded at a rate of 10 Hz by the BabyLog system [Dove *et al.* 1990, Ford *et al.* 1992, Macey 1998]. It is important for the system to be sealed, as it is the pressure change that generates the electrical signal. Sometimes infants breath *paradoxically*, where the infant's abdomen wall moves out of phase from the thoracic wall movement and breath inhalation [Haidmayer *et al.* 1980]. In this case, the Graseby signal is inverted in relation to inspiration and expiration. However, paradoxical breathing during quiet sleep is rare. Gross body movement causes noise in the signal, but cardiac movement is mostly filtered out. If gross movements cease, the pressure will gradually equalise to room pressure due to inevitable leaks in the equipment. Therefore, if the baby stops breathing, for instance, the signal will decay back to its mid-point due to the leaking of air from the system. In this thesis, unless stated otherwise, a rising signal indicates an exhalation whilst a decreasing signal indicates an inhalation.

Breathing signals from infants with both a low and a high risk of SIDS are studied in this thesis. An infant is classified as having a low risk of SIDS if the following criteria are met: non-smoking household, predominantly breastfed, slept on their backs or their

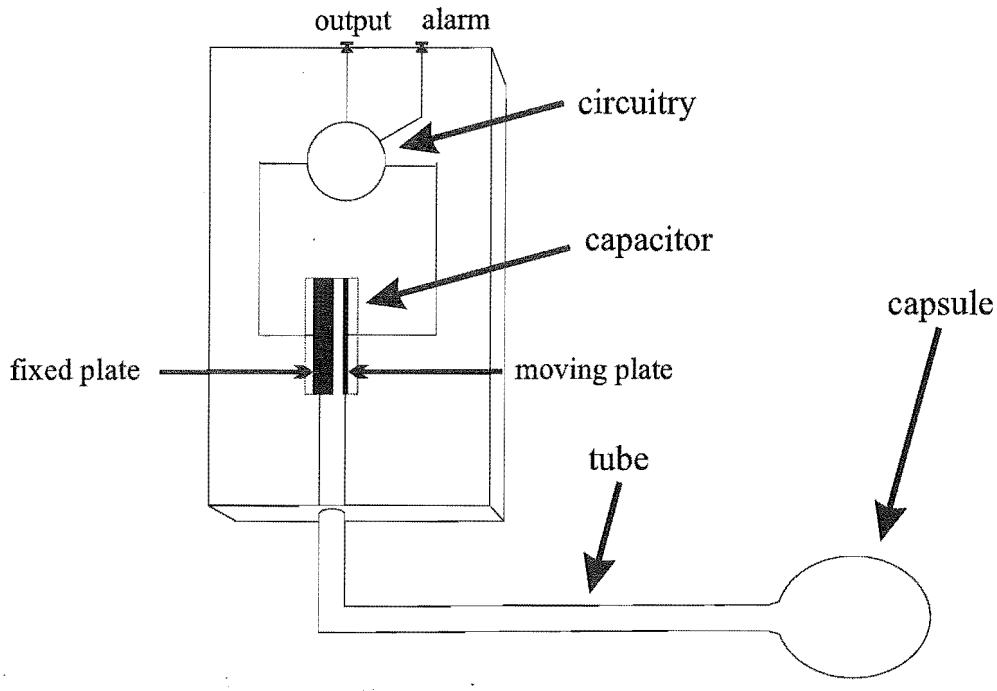


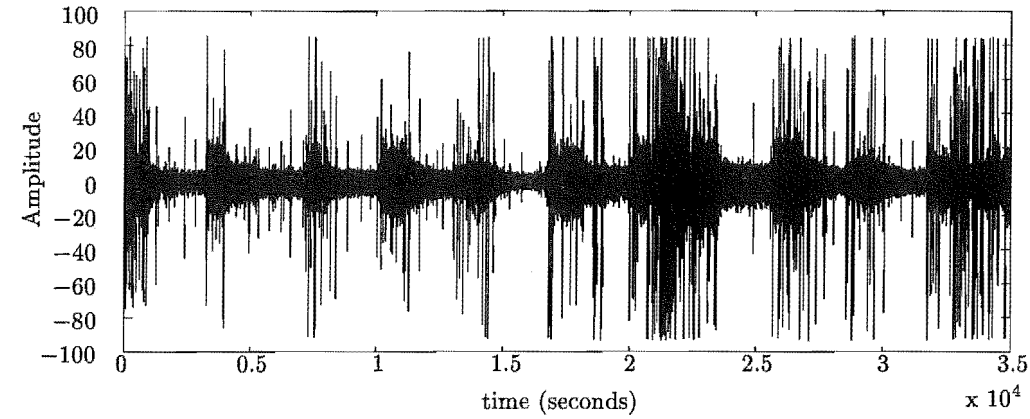
Figure 1.4 Schematic of the Graseby instrument used for detecting breathing.

sides, parents living together, parents older than 20 years old, non-Maori[1], not the first born in the family, no siblings that have died of SIDS, full term delivery, no serious illness or history of breathing problems, and a birth weight between 3000 and 4300 grams [Tappin *et al.* 1997]. A full discussion of these criteria is described by Tappin *et al.* [1997]. For this thesis, the high risk infants had all been admitted to hospital having suffered an ALTE [Ford *et al.* 1992].

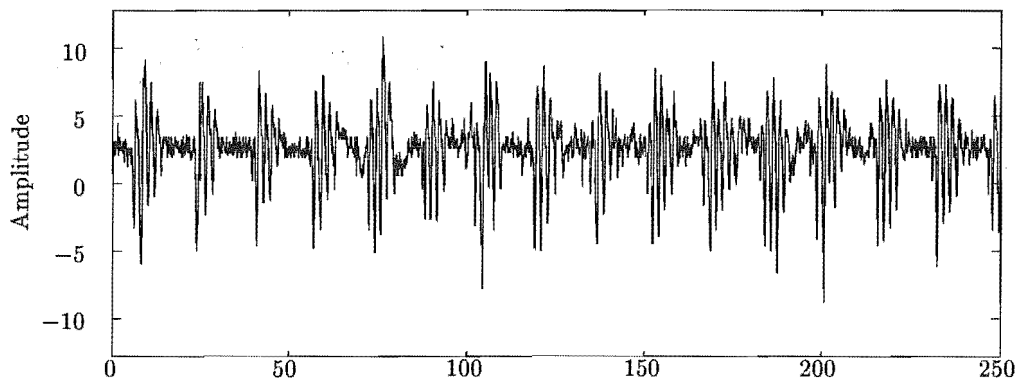
A second set of infants are studied in this thesis. These infants were part of a larger study held in Great Britain [Richards *et al.* 1984]. Respiration and cardiac signals were recorded for these infants. One group of these infants later succumbed to SIDS. The second group from these data, the controls for the infants who later succumbed to SIDS, were matched for age, gender and birth weight.

There are a myriad of patterns apparent in breathing signals of sleeping infants. These patterns occur over various time frames. On a large time frame (for example, a whole night of breathing as depicted in Figure 1.5(a)), the breathing is related to the type of sleep [Harper *et al.* 1987, Tappin *et al.* 1996a]. Sleep can be broadly classified as either Rapid Eye Movement (REM) sleep when high brain activity is noted, or quiet sleep when brain activity is low. During REM sleep, the infant's breathing is heavy and somewhat erratic in breath length and breath amplitude. The breathing during quiet sleep is more regular in breath length and breath amplitude. However, on a smaller time frame (such as a few minutes of breathing as shown in Figure 1.5(b)), when viewing the quiet sleep, other patterns emerge. There is a section of the signal where the amplitude

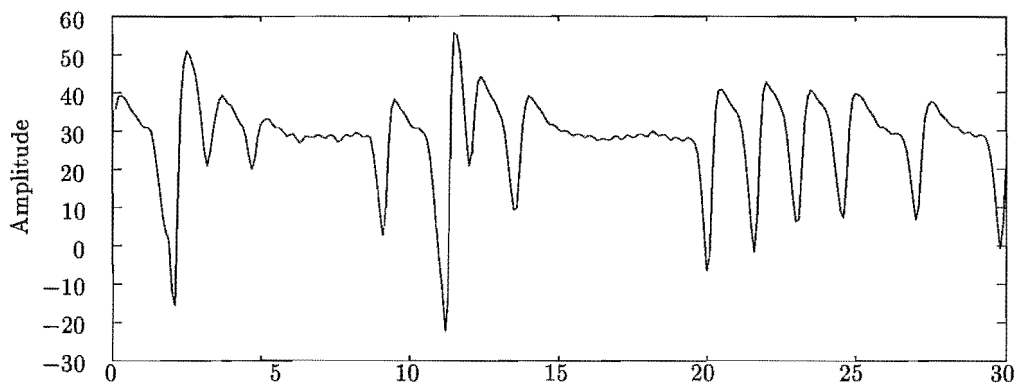
[1]Maori are the indigenous people of New Zealand.



(a) Breathing signal of an infant over night (9 hours, 43 minutes, 20 seconds).



(b) Breathing signal of an infant (4 minutes, 10 seconds).



(c) Breathing signal of an infant (30 seconds).

Figure 1.5 Three different views of breathing obtained from a sleeping infant.

of the signal has an envelope that is varying periodically. This pattern is described in this thesis as *amplitude modulation*. On a smaller time frame again (such as 30 seconds of breathing Figure 1.5(c)), periods where the infant stops breathing altogether, known as *apnoea*, are evident [Gibson 1996a]. Apnoea are commonly detected as part of hospital sleep study systems [Stein *et al.* 1979, Guilleminault *et al.* 1981, Haidmayer *et al.* 1982a, Hunt *et al.* 1985, Kelly *et al.* 1986, Oren *et al.* 1986, Lee *et al.* 1987, Kahn *et al.* 1992, Macey *et al.* 1995, Ford *et al.* 1996, Gibson 1996a, Tappin *et al.* 1996b]. A succession of apnoea at regular intervals is known as periodic breathing, such that the infant is only breathing for about half the time [Kempe *et al.* 1974, Gordon *et al.* 1984]. Periodic breathing could be considered to belong to the more general class of breathing where the normal breathing rate of the infant is modulated by a constant frequency, that is, amplitude modulation [Preiss *et al.* 1975, Rantonen *et al.* 1998]. This thesis investigates the use of wavelet transforms for analysing amplitude modulation of breathing signals.

1.3 THE ANALYSIS OF ULTRASOUND IMAGES

The second form of analysis considered in this thesis is of two-dimensional ultrasound images. This section introduces the approach taken in this thesis to image analysis. Ultrasound images and their features are also described in more detail.

In general, the aim of analysis in two-dimensions is to isolate features such as edges, regions of homogenous material, and textures. In many cases it is desirable to enhance images by making these features easier to see and interpret. Even though studies into two-dimensional feature extraction have been extensive, radiographers still prefer to trust their own judgment in examining raw data available from medical scans. This preference is due mainly to problems with isolating noise and artefacts from data in the scans [Anderson 1994]. Hence, there is a need for analyses that accurately and reliably enhance features and suppress noise and other artefacts while avoiding introducing new artefacts.

In this thesis, wavelet analysis is applied to ultrasound images with the intention of enhancing the images. Wavelets in two-dimensions can be classified as either separable or non-separable. Useful results using two-dimensional separable wavelets have been produced, especially in the arenas of de-noising, object detection and texture analysis [Healy, Jr. and Weaver 1992, Böröczky *et al.* 1995, Healy, Jr. *et al.* 1995, Laine *et al.* 1995, Richardson, Jr. 1995, Manjunath and Ma 1996, Hilton and Ogden 1997, Strickland and Hahn 1997]. Recently, results of texture analysis using non-separable wavelets has been reported [Mojsilović *et al.* 1997, Wang *et al.* 1998, Nicolier *et al.* 1999], however, overall there has been little reported use of non-separable wavelets in the literature. In this thesis, both separable and non-separable wavelets will be used for image enhancement.

Several image enhancement techniques applied in the wavelet domain under the broad categories of noise reduction and contrast enhancement are investigated and compared. Comparison of the outcomes of different enhancement techniques is difficult to perform, so an image quality equation is developed with the intention of assisting in the image comparison.

The two-dimensional medical images considered in this thesis are *B-scan ultrasound images* [McDicken 1991]. Ultrasound is widely used as an imaging tool for soft tissue applications. It is considered to be the safest medical imaging modality as the effects are negligible if operational guidelines are followed. Also, unlike X-rays, there are no cumulative negative effects from ultrasound.

B-scan ultrasound images are obtained by transmitting a beam of sound waves, typically at a frequency of about 5 MHz, and recording the echoes received from objects in the body. The important features in these images are regions of homogenous material and their edges. Homogenous material may be classified as skin, bone, fat, tumors, organs, and so on. Ultrasound is effective because each type of material in the body has different attenuation and reflective properties, which means that the resulting echo back to the receiver is different for each material. An echo is produced when a change in acoustical impedance is encountered by the ultrasound beam. Acoustical impedance, Z, is given by the ratio of applied pressure, p, to the resultant particle velocity in the material, u, that is $Z = \frac{p}{u}$. If it assumed that the material is a weak absorber then acoustical impedance can be approximated by $Z = \rho c$, where ρ is the density of the material, and $c = \sqrt{\frac{\kappa}{\rho}}$ is the speed of the ultrasound in the material, and where κ is the compressibility of the material. Different materials in the body have different densities and compressibilities and therefore, when going from one material to another, a change in acoustical impedance is encountered, $\Delta Z = \rho_1 c_1 - \rho_2 c_2$. The echo size is related to the magnitude of the change in acoustical impedance [McDicken 1991].

An image is created by displaying lines of echoes that are received in a plane with an intensity proportional to the echo size. Each line is collected very rapidly and whole images are typically produced at a rate of 25 images per second. The image capture is facilitated by the speed of the ultrasound waves in soft tissue, at an average of $c = 1540$ m/s [McDicken 1991].

The ultrasound images used in this thesis were obtained using an Acuson scanner at Christchurch Hospital. A frame grabber was used to capture a series of images and store them on computer [Surman *et al.* 1993]. The images are of a phantom of the female genital tract, provided by the company Schering. A sketch of the interior of the phantom is shown in Figure 1.6.

Ultrasound images are prone to artefacts. Artefacts in an ultrasound echo image are patterns of echoes that do not correspond to the structure being scanned. There are several different types and causes of artefacts, but only the types of artefacts encoun-

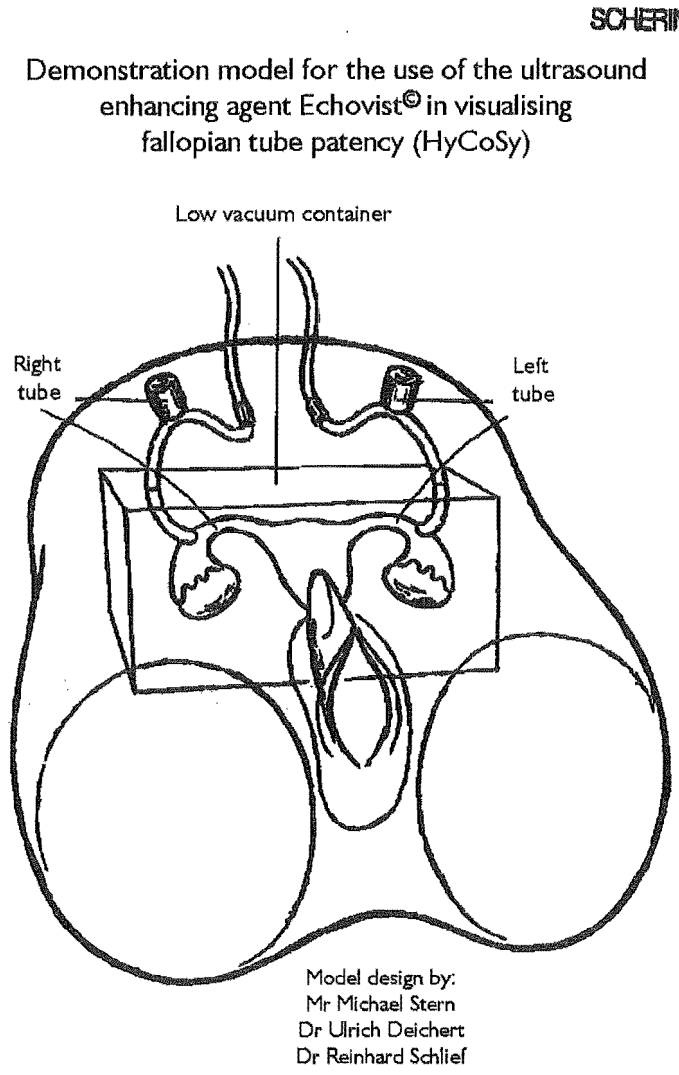


Figure 1.6 Sketch of phantom.

tered in the images considered in this thesis are discussed here. For a more complete discussion of artefacts, the reader is referred to McDicken [1991]. The artefacts considered here are shadowing, speckle, and beam widening.

Shadowing occurs when an object of high reflectivity causes the ultrasound beam to be reflected rather than transmitted through the object. In addition, shadowing is caused by a strongly absorbing layer. Shadowing results in a dark region, or shadow, to appear in the image behind the reflecting or absorbing layer.

Speckle occurs when the reflected echo is scattered. Scattered echoes from neighbouring parts of the object, or neighbouring objects, interfere with each other and distort the signal. The result is bright and dark "speckles" in the image, corresponding to the reinforcement and cancellation of the echoes.

Beam widening occurs outside the *focal zone*. The focal zone is the area of the image which is in focus. Inside the focal zone, point targets appear highly localised in

the image. The ultrasound beam is not ideal in that it is not narrow and uniform along its entire length, but instead, spreads outside the focal zone, making a point target appear as a short line in the image, defined as beam widening.

Artefacts can be identified in the image shown in Figure 1.7. This image is one of a series of images taken of the phantom of the female genital tract in Figure 1.6. The image was scanned from the left and shows the uterine cavity and fallopian tube as bright spots in the middle left and bottom centre of the image, respectively. Shadowing is evident to the right of the uterine cavity in the centre of the image. Speckle can be seen on the left of the image and beam widening to the right of the image. The location of these features are outlined in Figure 1.8

At present, human experts are still more adept at recognising features in ultrasound images than automatic techniques. However, eliminating, or at least reducing, the effects of artefacts in ultrasound images allows the images to be analysed more easily. This thesis uses two-dimensional wavelets with additional analysis in the wavelet domain to further improve the analysis, working towards the goal of automated analysis and enhancement of two-dimensional medical images.

1.4 SUMMARY

The focus of this thesis is the analysis of medical signals using wavelets. In this chapter, an introduction to wavelets is presented, as well as an introduction to the two types of medical signals that wavelets are applied to in this thesis: infant breathing signals and ultrasound images.

Wavelets and techniques in the wavelet domain are developed to analyse amplitude modulation of breathing from infants recorded during sleep. Amplitude modulation of breathing signals has not had much attention in the literature, mainly due to a lack of appropriate analysis tools. Only recently, one study included spectral analysis on regular breathing with amplitude modulation [Rantonen *et al.* 1998]. The aim of analysing these data with wavelets is to quantify differences in patterns between infants who succumb to Sudden Infant Death Syndrome (SIDS), or who are at high risk for SIDS and infants at low risk for SIDS. These data could provide information about the physiological control systems and possible deficiencies in the SIDS infants. Confirming differences may lead to improved prediction of the syndrome and, ultimately, to prevention of the syndrome.

Wavelets are also applied to a second type of medical signal, ultrasound images. These images are of poor quality with inherent artefacts. The images are difficult to enhance with objects in the image often appearing at similar intensity to the background noise level. The ultimate aim is for ultrasound images to be automatically enhanced which would allow radiologists and doctors to give better diagnoses. In this thesis, the aim is to improve the quality of the images and reduce the effects of the artefacts.

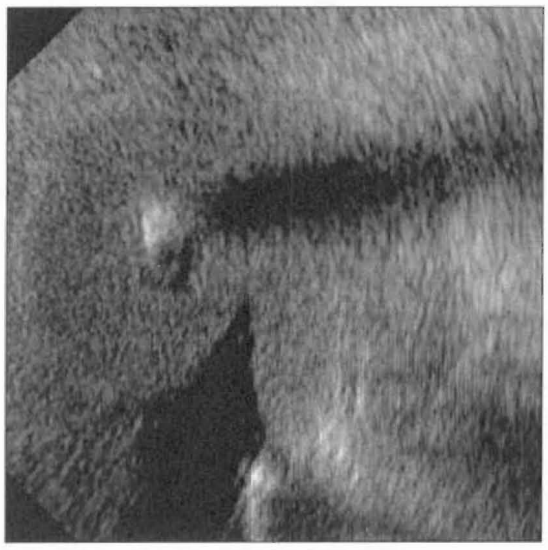


Figure 1.7 Ultrasound image of female genital tract.

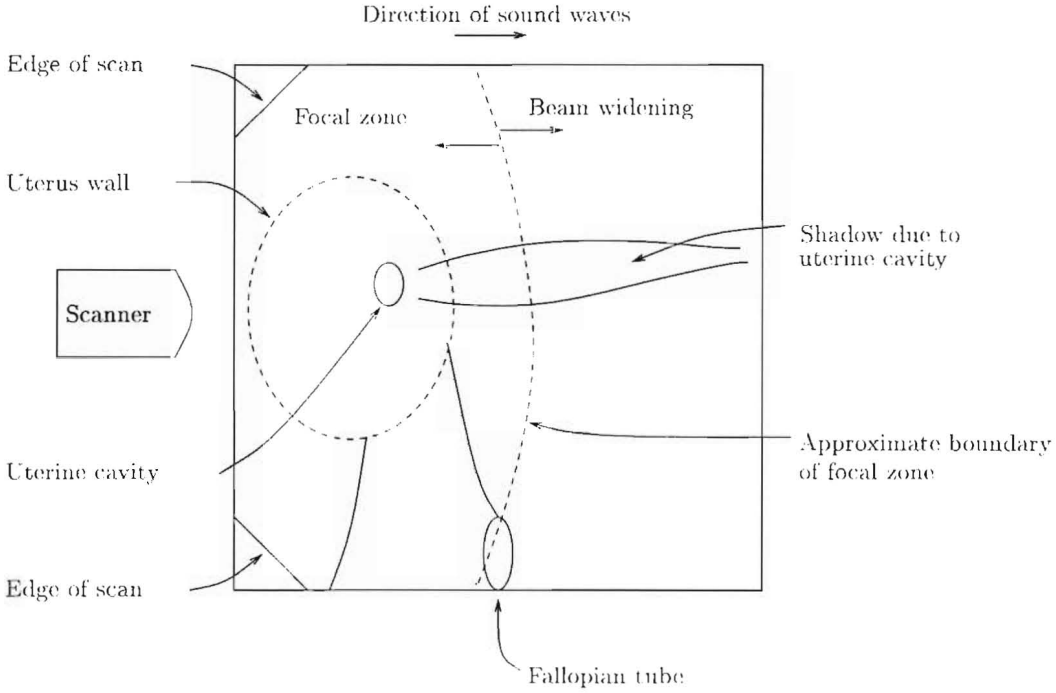


Figure 1.8 Line drawing of original ultrasound image.

Chapter 2

DEVELOPMENT OF WAVELETS

The theory of wavelet transforms is an amalgamation of theories from several disciplines. Historically, the three main contributions have been from the fields of mathematics, signal processing, and sub-band coding. Mathematically, the history of wavelets can be charted from Haar, with his work from the early 1900's based on the work of Fourier in the early 1800's and Du Bois-Reymond in the middle 1800's [Meyer 1993]. Since then, several lines of investigation have been drawn together to form the unified theory of wavelets. The second main contribution to the development of wavelets, which also dates back to Fourier, was from the field of signal processing. The approach was through a modification of the Fourier Transform, the Short Time Fourier Transform (STFT) [Rioul and Vetterli 1991, Bentley and McDonnell 1994]. The third main contribution was from the field of sub-band coding. Sub-band coding transforms a signal for storage or transmission and allows lossless reconstruction of the signal using quadrature mirror filters (QMFs) [Rioul and Vetterli 1991, Meyer 1993, Vetterli and Kovačević 1995]. Wavelets can be implemented and indeed constructed using quadrature mirror filters. This chapter describes the history and development of wavelets from these three perspectives. The concepts introduced in this chapter are developed in the following the chapter.

2.1 MATHEMATICAL HISTORY OF WAVELETS

This section introduces the concepts that lead to the mathematical underpinnings of wavelets. It is beyond the scope of this thesis to develop the entire mathematical background and the interested reader is referred to Meyer [1993], who is often referred to in this section, as opposed to the original mathematical papers.

In 1807, Joseph Fourier wrote that a 2π-periodic function could be represented by the sum of its Fourier Series, as it was to become known [Meyer 1993]. Prior to Fourier's assertion, polynomials were the only way to represent functions and the universes or spaces those functions described. By representing a function with its Fourier Series, Fourier had discovered a new functional universe. Fourier assumed that

2π-periodic functions converged everywhere. However, in 1878, Paul Du Bois-Reymond discovered a 2π-periodic function that diverged at a given point [Meyer 1993]. This discovery led to three options for mathematicians: change the definition of a function; change the definition of convergence; or, find other orthonormal systems that included the conditions discovered by Du Bois-Reymond. Haar [Meyer 1993] followed this last option and in 1909 reported what has become known as the Haar basis. A basis of a functional space is a set of independent functional elements that, when dilated and translated, span the functional space. Therefore, a set of basis functions, when combined as a weighted sum, can be used to construct any given signal from the space the basis spans [Bentley and McDonnell 1994]. The Haar basis is completely described by a piecewise function. As the limit for the number of Haar basis elements summed together is taken to infinity, this piecewise function can be used to construct any signal in the space of measurable, square integrable n-dimensional functions, that is, the $L^2(\mathcal{R}^n)$ functional universe, where $\mathcal{R}^n$, the Euclidean space, is the set of real numbers in n-dimensions. However, the piecewise nature of the Haar basis has limitations in the finite case when approximating continuous functions.

Faber and Schauder worked on the Haar basis between 1910 and 1920 [Meyer 1993]. A major improvement was the use of continuous functions as the basis. This new basis was called the Schauder basis. The continuity of the elements contrasts with the piecewise elements of the Haar basis. Using continuous functions is an improvement because many functions being approximated are at least continuous, and in most cases are continuous in the first derivative also. Approximation of a function that is continuous by a basis that is continuous does not have the same limitations as approximation by a piecewise basis. This feature of the Schauder basis points to one of the desirable features of wavelets: they are at least continuous.

Faber and Schauder's study of functions was expanded by four separate groups in the 1930's. Paul Lévy, with his work on Brownian motion, showed in the 1930's that the Schauder Basis was superior to Fourier analysis for studying local regularity properties [Meyer 1993]. Regularity is related to the continuity of a function. Both regularity and continuity may change depending on the location in a function, and hence, the Schauder basis is suited to studying localised properties. The ability to study local regularity properties is another desirable feature of wavelets.

Also in the 1930's, Littlewood and Paley were studying the localised energy of a function [Meyer 1993]. They succeeded in this for the one-dimensional periodic case by using *dyadic blocks*. Dyadic blocks are essentially defined portions of the Fourier Series, each spanning an octave in the frequency range. The notion of dyadic blocks is used in wavelet theory. Antoni Zygmond, and the mathematicians in his group, extended Littlewood and Paley's work to n-dimensional Euclidean space and produced the Littlewood-Paley-Stein function, which led directly to the concept of the "mother wavelet" [Meyer 1993].

The mother wavelet is a basis function with the following properties. The mother wavelet is defined on the Euclidean space, $\mathcal{R}^n$, as a rapidly decreasing function of time or space, t. Both the wavelet, $\psi(t)$, and its Fourier Transform, $\Psi(f)$, are infinitely differentiable. An infinitely differentiable function is infinitely continuous. The power of the Fourier Transform of the mother wavelet is bounded and has a magnitude of 1 within those bounds and a magnitude of 0 outside them. These features of the mother wavelet imply that the Littlewood-Paley-Stein function conserves energy and provides a method for analysing the function of interest, $x(t)$. The sequence of operators given by the Littlewood-Paley-Stein function constitutes a bank of band-pass filters, orientated on frequency intervals covering approximately one octave. This work led to an effective algorithm for numerical image processing through work by Marr and Mallat [Mallat 1989].

Lusin was working in the 1930's with complex Hardy spaces, denoted by $\mathcal{H}^p(\mathcal{R})$ [Meyer 1993]. He was interested in developing elementary functions of $\mathcal{H}^p(\mathcal{R})$, which he called "atoms" or "basis elements" to use in the analysis and synthesis of functions in $\mathcal{H}^p(\mathcal{R})$. Weiss and Coifman picked up Lusin's work in the 1960's, continuing to pursue the idea of atomic decompositions being used to make up a function. In the 1960's, Calderón, and independently Grossmann and Morlet, discovered a decomposition of the identity operator that led directly to wavelets [Meyer 1993].

The fourth study in the 1930's based on Faber and Schauder's work was that done by Franklin. Franklin developed an orthonormal basis from the Schauder basis by using the Gram-Schmidt process [Meyer 1993]. This development resulted in a basis that had the advantages of both the Haar and Schauder bases. However, the disadvantage was that the simple algorithmic structure for producing either of the other two bases had been lost. In 1963, Zbigniew Ciesielski revived Franklin's work when he discovered the existence of two parameters giving a wavelet-like function that is known as a Lipschitz function with an exponential decay [Meyer 1993].

In 1980, Strömberg [Meyer 1993] discovered, when examining real Hardy spaces, $\mathcal{H}^p(\mathcal{R})$, an orthonormal wavelet basis as an unconditional basis of $\mathcal{H}^p(\mathcal{R})$. This was very similar to Franklin's work.

These ideas—that a function can be represented by weighted "atomic decompositions," that the atomic decompositions can be continuous (and even infinitely continuous), that they can be used to analyse local regularity properties or energy of a function, and that the atomic decompositions could form an orthonormal basis—together form a definition of a wavelet basis that is capable of analysing function spaces other than $L^2(\mathcal{R}^n)$, the space of measurable, square integrable n-dimensional functions.

These conditions for basis functions (regularity, decay at infinity, and some condition on the number of vanishing moments, which is related to the continuity of a function) led to work by Mallat [1989]. Mallat combined this work with the theory of

quadrature mirror filters and pyramid algorithms to further describe wavelets. Mallat's work allowed Ingrid Daubechies [1992] to extend Haar's original thesis and produce compact orthogonal wavelets.

2.2 SIGNAL PROCESSING HISTORY OF WAVELETS

The second approach to the development of wavelet theory is from a signal processing viewpoint. This approach originates from the Fourier Transform, which is modified for some applications to create the Short Time Fourier Transform (STFT). The STFT is again modified, resulting in the Continuous Wavelet Transform [Rioul and Vetterli 1991, Bentley and McDonnell 1994]. This section describes the development of the Continuous Wavelet Transform by tracing its development through the Fourier Transform and the Short Time Fourier Transform.

2.2.1 The Fourier Transform

The Fourier Transform is most commonly used as a tool for determining the frequency spectrum of a signal. Generally, signals are measured either in the time domain or in the spatial domain, with units denoted by t. The Fourier Transform may be applied to either of these domains, transforming the signal, $x(t)$, into the frequency domain, units f, giving the Fourier Transform of the signal, $X(f)$, as:

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-2i\pi ft}dt. \tag{2.1}$$

The Fourier Transform is useful for analysing stationary signals. However, most real world signals are non-stationary; they contain transient activity that results in a changing frequency spectrum over the length of the signal. To analyse transient signals, not only the frequency information but also the locational or temporal information is important. If the spectral information changes within the section of signal being studied, then this information is lost when looking at the magnitude plots of the Fourier Transform. Two real signals which have different spectral properties could have similar Fourier Transforms. For instance, the magnitude of the Fourier Transforms in Figures 2.1(a) and 2.1(b) are virtually identical but the original signals in Figures 2.1(c) and 2.1(d) are quite different. The signal in Figure 2.1(c), with Fourier Transform in Figure 2.1(a), is made up of three continuous sine waves, and, therefore, has a constant spectrum. The signal in Figure 2.1(d), with Fourier Transform in Figure 2.1(b), consists of three waves of the same frequencies as Figure 2.1(c), but that are spatially distinct, therefore, having an evolving spectrum across the signal. The spatial information regarding the differences is lost in the magnitude plots of the Fourier Transform.

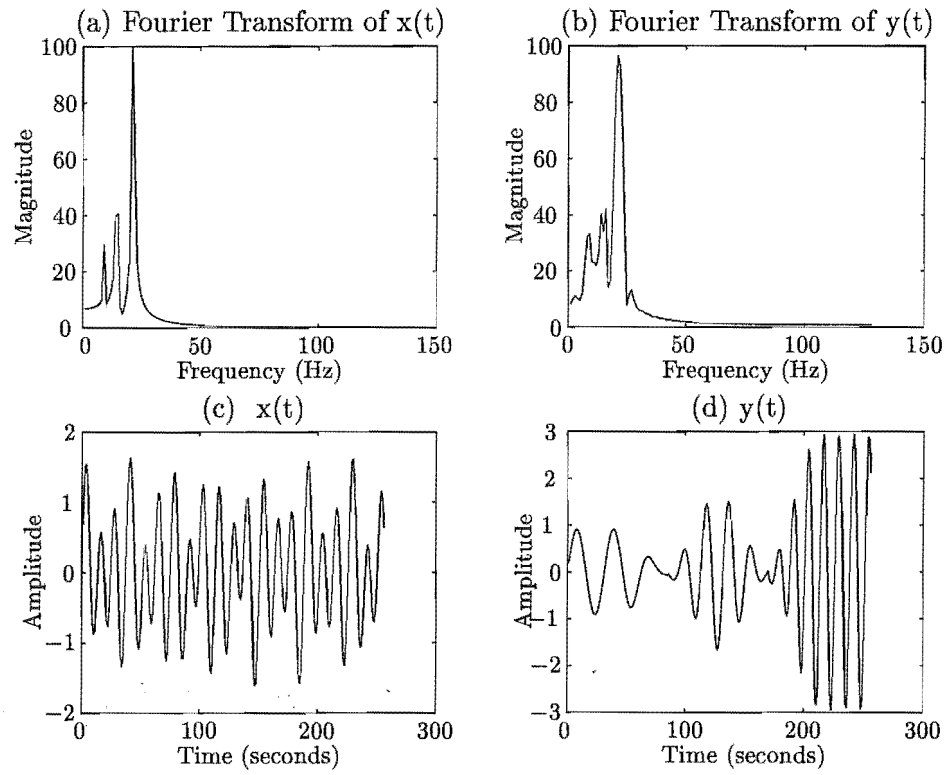


Figure 2.1 (a) and (b) are the magnitudes of Fourier Transforms of the signals directly below. (c) is a stationary signal. (d) has transient components. The magnitudes of their Fourier Transforms are similar.

Thus, the Fourier Transform is useful for discovering the frequency content of a signal, but it is less useful if the information required includes the location of the frequency content in the signal. Therefore, another approach was investigated. This approach was similar to Lévy's work with Schauder bases, and to Littlewood and Paley's investigation into the localised energy of a function [Meyer 1993].

2.2.2 The Short Time Fourier Transform

In 1946, Gabor [1946] introduced a transform that used a windowed analysing function to investigate the localised energy of a signal. This transform was the Short Time Fourier Transform (STFT). The STFT takes windows, $g(t)$, of the signal, $x(t)$, in the time domain centered at varying time locations, τ, and representing the frequencies within each window. The STFT is defined as follows, with * denoting the complex conjugate:

$$X(\tau, f) = \int x(t)g^*(t-\tau)e^{-2i\pi ft}dt. \tag{2.2}$$

If the windows $g(t)$ are of a short enough duration, then the signal within the window may be considered stationary. Therefore, performing the Fourier Transform within the window will give an accurate reflection of the frequency components across the whole

window, whilst the position of the window gives the locational information.

The accuracy of this approach in determining the localised energy of the signal is limited by the Heisenburg Uncertainty Principle, which states that the product of the transform resolutions in the spatial domain, Δt, and in the frequency domain, Δf, is always greater than, or at best equal to, a constant [Rioul and Vetterli 1991, Bentley and McDonnell 1994, Lee and Yamamoto 1994], as described by the following equation:

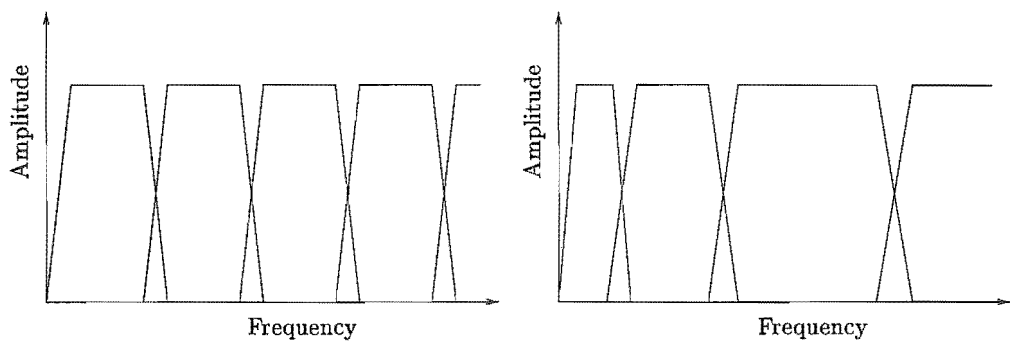
$$\Delta t \Delta f \geq \frac{1}{4\pi}. \tag{2.3}$$

Δt and Δf can be approximated using the root-mean-square departures from the centroids, i.e. variance, of the energy distribution and energy spectrum, respectively [Bracewell 1986]. In practical terms, a long spatial window is required for high resolution in the frequency domain, but this results in poor localisation in the spatial domain. By optimising the window length, the energy localisation can be determined within the limitations given by the Heisenburg Uncertainty Principle.

The STFT has a constant bandwidth over all frequencies since the windows in the spatial domain are of a constant length. The frequency response of the transform is represented by constant bandwidths as shown in Figure 2.2(a).

The Short Time Fourier Transform is represented in the time-frequency plane as even *tiles* (see Figure 2.3 (a)). Tiles can be thought of as the area represented by each of the basis elements in terms of time and frequency. Each tile represents a certain resolution of the transform in the time and frequency domains. In the Short Time Fourier Transform, the tile sizes are limited in both dimensions by choosing a constant time window length and therefore bandwidth for the transform.

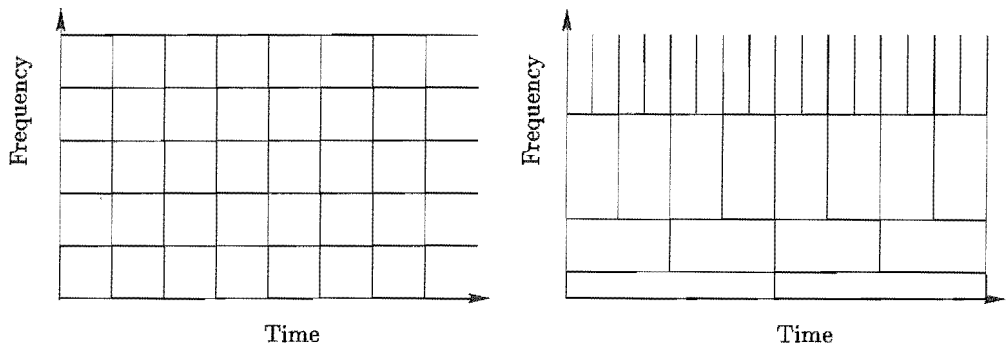
Unlike the Short Time Fourier Transform, the Wavelet Transform exploits the



(a) Short Time Fourier Transform frequency response. Each window is represented by a constant bandwidth.

(b) Wavelet Transform frequency response. Each window is represented by a constant relative bandwidth.

Figure 2.2 Short Time Fourier Transform (STFT) has a constant bandwidth compared with the Wavelet Transform (WT) which has a constant relative bandwidth.



(a) Short Time Fourier Transform frequency-time tiles.

(b) Wavelet Transform frequency-time tiles.

Figure 2.3 Time-Frequency tilings.

Heisenburg Uncertainty Principle by varying the window length in time and frequency at different *scales*. The scales do not relate exactly to frequency, but typically cover an octave within each window of analysis. The frequency is varied so that the relative bandwidth is kept constant, i.e. $\frac{\Delta f}{f} = c$, as in Figure 2.2(b). Even though both time and frequency dimensions change for each scale in the time-frequency plane, the tiles have equal area (see Figure 2.3 (b)).

The following section introduces a mathematical definition for scales.

2.2.3 Multiresolutional Analysis

Wavelet analysis may be considered as a class of multiresolutional analysis. Understanding multiresolutional analysis provides a framework upon which wavelet theory in Section 2.3 and Chapter 3 is built.

Multiresolutional analysis involves taking approximations of the signal at successively coarser resolutions, and finding the difference between the approximation at one resolution and the next coarser resolution. The approximations are written as a linear combination of a scaling function, and the differences can be written as a linear combination of a corresponding wavelet function [Mallat 1989, Daubechies 1992, Daubechies 1993]. The multiresolution approach is from the perspective of mathematical spaces, following Mallat [1989]. Three spaces representing the signal, the approximations of the signal, and the details of the signal are defined. These spaces are defined in terms of orthogonal bases and in terms of the orthogonality between spaces.

Wavelet theory applies to measurable, square integrable n-dimensional signals, i.e. to a signal, $x(t)$ within the vector space $L^2(\mathcal{R}^n)$. The signal, $x(t)$, has finite energy.

A linear operator, A_{2^j}, approximates the signal at resolution 2^j, where scale, $j \in \mathcal{Z}$. A space, V_{2^j}, is the set of all possible approximations of signals in $L^2(\mathcal{R}^n)$ at the

resolution 2^j. If A_{2^j} is an orthogonal projection operator on the vector space $V_{2^j} \subset L^2(\mathcal{R}^n)$, then the operator A_{2^j} is characterised by the following five properties.

1. The approximating signal, $A_{2^j}x(t)$, is taken as the signal that is most similar to $x(t)$, that is:

$$\|A_{2^j}x(t) - x(t)\| \leq \|y(t) - x(t)\|, \quad \forall y(t) \in V_{2^j}. \tag{2.4}$$

2. The approximation of the signal at resolution 2^{j+1} contains the information required to compute the signal at the smaller resolution of 2^j. This property is written as:

$$V_{2^j} \subset V_{2^{j+1}}, \quad \forall j \in \mathcal{Z}. \tag{2.5}$$

3. A_{2^j} is self-similar at all resolutions and so the spaces of approximation functions, V_{2^j}, are obtained by scaling each approximation function by the ratio of resolutions, that is:

$$x(t) \in V_{2^j} \Leftrightarrow x(2t) \in V_{2^{j+1}}, \quad \forall j \in \mathcal{Z}. \tag{2.6}$$

4. Translations of $x(t)$ and $A_{2^j}x(t)$ are given by the following general rules. The approximation $A_{2^j}x(t)$ of signal $x(t)$ is characterised by 2^j samples per unit length. This characterisation gives the Discrete Wavelet Transform, as described in Section 2.3. If $x(t)$ is translated by a length proportional to 2^{-j} then $A_{2^j}x(t)$ is translated by the same amount and is characterised by the same samples which have been translated. These rules are expressed mathematically as:

   - Translations are a discrete characterisation such that:

$$\exists \text{ an isomorphism } \mathcal{I} \text{ from } V_1 \text{ onto } \mathcal{I}^2(\mathcal{Z}), \tag{2.7}$$

     where $\mathcal{I}^2(\mathcal{Z})$ is the vector space of all square-summable sequences. That is

$$\mathcal{I}^2(\mathcal{Z}) = \{(\alpha_l) : \sum_{l=-\infty}^{\infty} |\alpha_l|^2 < \infty\}, \quad l \epsilon \mathcal{Z}.$$

   - Translations of approximations are given by:

$$A_1 x_k(t) = A_1 x_k(t-k), \quad \forall\, k \in \mathcal{Z},, \tag{2.8}$$

     where $x_k(t) = x(t-k)$.

- Translation of samples is given by

$$\mathcal{I}(A_1 x(t)) = (\alpha_l), \quad l \epsilon \mathcal{Z} \Leftrightarrow \mathcal{I}(A_1 x_k(t)) = (\alpha_{l-k}), \quad l, k \epsilon \mathcal{Z} \tag{2.9}$$

5. At each resolution 2^j, some information about $x(t)$ is lost when calculating the approximation. As the resolution tends to infinity, more and more information is known about the signal and the approximation converges to the signal. Conversely, as the resolution approaches zero, less is known about the signal and the approximation converges to zero. This is represented in vector space as:

$$\lim_{j \to \infty} V_{2^j} = \bigcup_{j=-\infty}^{\infty} V_{2^j} \text{ is dense in } L^2(\mathcal{R}^n), \tag{2.10}$$

$$\lim_{j \to -\infty} V_{2^j} = \bigcap_{j=-\infty}^{\infty} V_{2^j} = 0. \tag{2.11}$$

A *multiresolution approximation* of $L^2(\mathcal{R}^n)$ is defined by a set of vector spaces, V_{2^j}, that satisfies the properties (2.5) to (2.11). The approximation of any $L^2(\mathcal{R}^n)$ function at resolution 2^j is given by the associated set of linear operators, A_{2^j}, satisfying properties (2.4) to (2.9).

To numerically characterise A_{2^j}, an orthonormal basis for V_{2^j} must be found. Such a basis may be found by dilating and translating a unique function, $\phi(t) \epsilon L^2(\mathcal{R}^n)$, which is known as the *scaling function.* If the dilation of the scaling function is set such that:

$$\phi_{2^j}(t) = 2^j \phi(2^j t), \ \forall j \epsilon \mathcal{Z}, \tag{2.12}$$

then an orthonormal basis of V_{2^j} is given by

$$(\sqrt{2^{-j}} \phi_{2^j}(t - 2^{-j} n)), \quad n \epsilon \mathcal{Z}. \tag{2.13}$$

The projection on V_{2^j} is produced by decomposing the signal, $x(t)$, on the orthonormal basis, $\phi(t)$, as given in Equation (2.13):

$$A_{2^j} x(t) = 2^{-j} \sum_{n=-\infty}^{\infty} \langle x(u), \phi_{2^j}(u - 2^{-j} n) \rangle \, \phi_{2^j}(t - 2^{-j} n), \quad \forall x(t) \epsilon L^2(\mathcal{R}^n). \tag{2.14}$$

The approximation of the signal, $x(t)$ at the resolution 2^j is characterised by a set of discrete approximations of $x(t)$ at the resolution 2^j, $A^d_{2^j} x(t)$, where the set is given by the inner products:

$$A^d_{2^j} x(t) = \langle x(u), \phi_{2^j}(u - 2^{-j} n) \rangle, \quad n \epsilon \mathcal{Z}. \tag{2.15}$$

Each inner product can also be represented as a convolution evaluated at a point $2^{-j}n$:

$$A_{2^j}^d x(t) = (x(u) \odot \phi_{2^j}(-u))(2^{-j}n), \quad n \in \mathcal{Z}, \tag{2.16}$$

which is essentially low-pass filtering of $x(t)$ followed by uniform sampling at a rate of 2^j. This equation is used in Section 2.3 and Chapter 3 to implement wavelet transforms.

Summarising so far, there exists a set of spaces that represent the orthogonal projection of approximations to the signal onto those spaces. For each resolution, there is part of the signal that is not represented by the approximation at the next finer resolution. This is called the detail of the signal at that resolution.

There is a set of spaces that represent the details of the signals at all resolutions, similar to the space representing the approximations to the signal. The spaces that represent the approximations and the details are orthogonal to each other. Approximation and detail pairs of spaces at a particular resolution can be combined to give the approximation space at the next finer resolution. The concepts associated with the details of the signal and corresponding spaces are expanded below.

The approximation to a signal, $A_{2^j}^d x(t)$, is necessarily a smoothing of the signal at a given resolution, 2^j. The difference of information between this resolution and the next resolution, 2^{j+1}, is called the detail signal at the resolution 2^j. The detail signal at resolution 2^j is given by the orthogonal projection of the original signal on the orthogonal complement of V_{2^j} in $V_{2^{j+1}}$. O_{2^j} is the orthogonal complement to V_{2^j},

$$O_{2^j} \oplus V_{2^j} = V_{2^{j+1}} \tag{2.17}$$

where $\oplus$ represents appending the two vector spaces together.

An orthonormal basis for O_{2^j} is constructed by dilating and translating a function $\psi(t)$. $\psi(t)$ is called the wavelet function and is described by the dilation of $\psi(t)$ by 2^j:

$$\psi_{2^j}(t) = 2^j \psi(2^j t), \quad \forall\, j \in \mathcal{Z}.$$

The orthonormal basis of O_{2^j} is given by

$$(\sqrt{2^{-j}}\psi_{2^j}(x - 2^{-j}n)), \quad n \in \mathcal{Z}. \tag{2.18}$$

Taking the projection on O_{2^j} by decomposing the signal, $x(t)$, on the orthonormal basis $\psi(t)$ leads to band-pass filtering of the signal. This result is used in Section 2.3 and Chapter 3 to implement orthogonal wavelet transforms.

This review of multiresolution analysis has concentrated on orthogonal wavelets and scaling functions. A second type of wavelets known as *biorthogonal wavelets* can also be developed, where the combination of the filter operations is a half-band filter. Biorthogonal wavelets and scaling functions lead to linear phase filters—a property

that is discussed in Chapter 3. In this case, two more spaces are introduced such that [Strang and Nguyen 1996]:

$$V_{2j} \oplus O_{2j} = V_{2j+1}, \text{ and } \tilde{V}_{2j} \oplus \tilde{O}_{2j} = \tilde{V}_{2j+1},$$

where

$$V_{2j} \perp \tilde{O}_{2j}, \text{ and } O_{2j} \perp \tilde{V}_{2j} \tag{2.19}$$

and $\perp$ means "orthogonal to."

Basis functions for $\tilde{V}_{2j}$ and $\tilde{O}_{2j}$ are found in a similar manner as V_{2j} and O_{2j}. Construction is achieved by dilating and translating functions $\tilde{\phi}(t)$ and $\tilde{\psi}(t)$ giving

$$\begin{aligned} &(\sqrt{2^{-j}}\tilde{\phi}_{2j}(x - 2^{-j}n)), \\ &\text{and } (\sqrt{2^{-j}}\tilde{\psi}_{2j}(x - 2^{-j}n)), \quad n \epsilon \mathcal{Z}. \end{aligned} \tag{2.20}$$

A dual wavelet basis is formed by ψ with ϕ and $\tilde{\psi}$ with $\tilde{\phi}$.

This section has developed a background to wavelet analysis through multiresolutional analysis theory. The theory explains how orthogonal or biorthogonal projections of the signal onto two or four different spaces, respectively, which are themselves orthogonal complements, can completely describe the signal at any given resolution.

2.2.4 The Continuous Wavelet Transform

The Continuous Wavelet Transform (CWT) is defined in this section. The CWT is calculated at different scales. The scales, usually denoted a, are related to each resolution 2^j from the last section. For each scale a, the wavelet $\psi(t)$ is a scaled version of itself as seen in Section 2.2.3:

$$\psi_a(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t}{a}\right). \tag{2.21}$$

Translations of the wavelet by τ are given by:

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-\tau}{a}\right). \tag{2.22}$$

Approximations of wavelets can be seen in Figure 2.4 (b). For comparison, approximations of Short Time Fourier Transform functions, $g(t - \tau)$, where the window length is constant, can be seen in Figure 2.4 (a).

For each dilation, or scale, a, and translation τ, the Wavelet Transform, $X(\tau, a)$, of

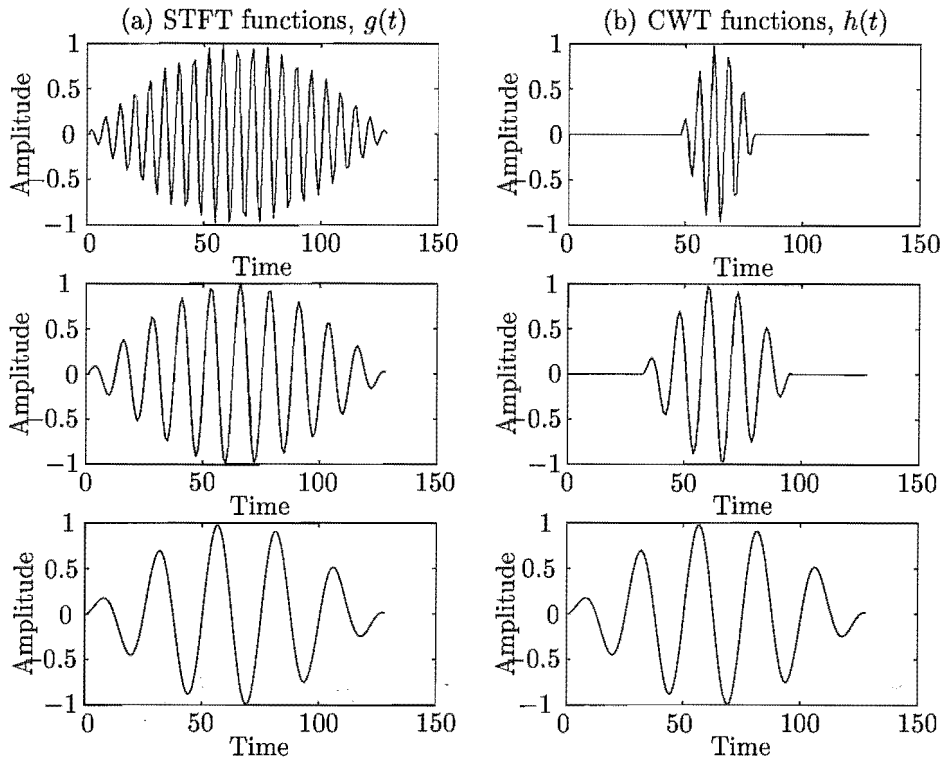


Figure 2.4 Approximations of (a) Short Time Fourier Transform functions and (b) Wavelet Transform functions for three different scales.

the signal, $x(t)$, indicates how closely the signal is represented by the wavelet, $\psi_{a,\tau}(t)$,

$$X(\tau, a) = \frac{1}{\sqrt{|a|}} \int x(t)\psi^* \left(\frac{t-\tau}{a}\right) dt, \tag{2.23}$$

i.e. the Wavelet Transform measures the similarity between the signal and basis functions which are the wavelets $\psi_{a,\tau}(t)$.

Synthesis, or reconstruction of the signal, consists of summing up all the orthogonal projections of the signal onto the wavelets:

$$x(t) = c \int\int_{a<0} \frac{1}{a^2} X(\tau, a)\psi_{a,\tau}(t)\, da\, , d\tau \tag{2.24}$$

where c is a constant that depends only on $\psi(t)$. If it is assumed that both the signal and wavelets are either real-valued or complex analytic, then only positive dilations, where $a > 0$, have to be taken into account [Rioul and Vetterli 1991]. Equation (2.24) is satisfied whenever $\psi(t)$ is of finite energy and band-pass. The condition that $\psi(t)$ be of finite energy and band-pass is known as the reconstruction condition.

A more precise definition of the reconstruction condition depends on the regularity of ψ. Regularity is related to differentiability [Vetterli and Kovačević 1995]. It can be described using the space of continuous functions, C^0, and the spaces of functions having n continuous derivatives, C^n, $n \in \mathcal{Z}$. For regularity defined on smaller intervals,

Lipschitz or Hölder exponents are used. A function $x(t)$ is called Lipshitz of order $\alpha, 0 < \alpha \leq 1$, if, for any t and some small ζ [Vetterli and Kovačević 1995],

$$|x(t) - x(t-\zeta)| \leq c|\zeta|^{\alpha}, \tag{2.25}$$

where c is a finite constant. Functions of higher Lipshitz orders, $r = n + \alpha$, can be found by replacing $x(t)$ with its nth derivative. Equation (2.25) and its corresponding higher orders define Hölder spaces of regularity and order r. Therefore, if $\psi(t)$ is assumed sufficiently regular, then the reconstruction condition is given by [Bentley and McDonnell 1994, Grossmann *et al.* 1987]:

$$\int \psi(t)dt = 0. \tag{2.26}$$

If the reconstruction condition is satisfied, then the Continuous Wavelet Transform preserves the energy of the signal, E_x, [Grossmann *et al.* 1987]:

$$\int\int \frac{1}{a^2}|X(\tau, a)|^2 d\tau da = E_x = \int |x(t)|^2 dt. \tag{2.27}$$

2.2.5 Representing the Continuous Wavelet Transform

The continuous wavelet transform, $X(\tau, a)$, of a signal is represented on a time-scale plane on a plot known as a *scalogram*, as shown in Figure 1.2(b). Usually, the time dimension is in the horizontal direction and the scale dimension is in the vertical direction. Low scales are represented below high scales. There is an inverse relationship between scale and frequency. Therefore, high frequencies are mapped below low frequencies, which is the inverse of the convention for frequency [Grossmann *et al.* 1987]. In some cases, the scale is represented with small scales above larger scales in order to comply with the convention for frequency [Grossmann *et al.* 1987]. Complex valued transforms are often represented by the phase and modulus separately [Grossmann *et al.* 1987, Rioul and Vetterli 1991, Lee and Yamamoto 1994]. Since the Continuous Wavelet Transform preserves the energy of the signal, as in Equation (2.27), the scalogram can be defined as the squared modulus of the Continuous Wavelet Transform [Rioul and Vetterli 1991]. Others define the scalogram using the modulus of the Continuous Wavelet Transform only [Grossmann *et al.* 1987, Lee and Yamamoto 1994]. For some applications, such as analysing sound signals in the audible range, a large range of scales is used. To map a large scale range on a scalogram, the log of the scale is taken [Grossmann *et al.* 1987, Lee and Yamamoto 1994].

In this thesis, time is always plotted in the horizontal direction and scales in the vertical direction. Scales are plotted with low scales below higher scales. Therefore, high frequencies are represented at the bottom of the plot, and lower frequencies at the top of the plot.

2.3 SUB-BAND CODING HISTORY OF WAVELETS

The third approach to developing the theory of wavelets is through sub-band coding [Rioul and Vetterli 1991, Kovačević and Vetterli 1992, Vetterli and Kovačević 1995, Strang and Nguyen 1996]. Sub-band coding developed out of the need to encode signals to allow their efficient transmission or storage and subsequent recovery. Initial applications were in speech [Croisier *et al.* 1976, Esteban and Galand 1977] and image coding [Mallat 1989]. Sub-band coding systems were first introduced by Croisier, Esteban and Galand, [Croisier *et al.* 1976, Esteban and Galand 1977] in 1976 for speech compression. Galand saw two advantages of introducing sub-band coding: firstly, the algorithm was simple and, secondly, the effects of quantisation noise, as perceived by the receiver, were reduced compared with other types of coding. Quantisation noise was less apparent as quantisation was done inside each sub-band and the signal therefore tended to mask the quantisation noise [Meyer 1993]. Although Galand applied sub-band coding to the transmission of speech, the argument regarding the apparent reduction of quantisation noise can be used when considering analysis of a signal in each sub-band, as this thesis does. The same argument has been used for digital image processing [Meyer 1993].

In sub-band coding, the signal is encoded then transmitted or stored prior to being decoded. For signal processing applications, the encoding step is equivalent to an analysis step, and the decoding step is equivalent to reconstruction of the signal, or synthesis (see Figure 2.5).

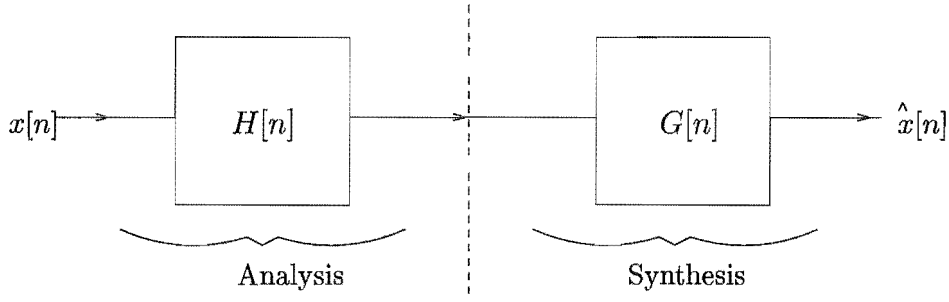


Figure 2.5 An analysis and synthesis filter bank, with a signal, $x[n]$, where n is the sample number. $H[n]$ is the analysis function and $G[n]$ is the synthesis function. $\hat{x}[n]$ is the reconstructed signal.

The aim of sub-band analysis and synthesis systems is perfect reconstruction, that is, $\hat{x}[n] = x[n]$. For coding applications, the reconstructed signal should be identical to the original signal. For signal processing applications, some additional processing may occur between the analysis and synthesis steps. However, even in signal processing, it is desirable for the analysis and synthesis stages of the process to be a perfect reconstruction system. In analysis and synthesis systems there are several problems that prevent perfect reconstruction. The problems encountered are aliasing, shift variance and distortions in the amplitude and phase [Meyer 1993]. Aliasing cancellation was achieved with the advent of quadrature mirror filters [Vetterli and LeGall 1989]. These

filters are also shift-invariant in the absence of coding loss. If amplitude and phase distortion are also eliminated, then the original signal is recovered and the sub-band coding system is called a perfect reconstruction system [Vetterli and LeGall 1989].

2.3.1 Sub-band Coding

This section introduces the theory of sub-band coding.

Take a signal, $x[n]$, with samples denoted by integers, n, and with finite energy

$$\sum_{n=-\infty}^{\infty} x[n]^2 < \infty, \quad n \in \mathcal{Z}. \tag{2.28}$$

Sub-band coding is performed by filtering the signal through M frequency channels. Each channel is associated with an M^{th} of the frequency band of the original signal. By the Nyquist theorem, since each signal has only an M^{th} of the frequency band of the original, the filtered signal can be re-sampled so that only every M^{th} sample remains. This is called *downsampling by* M, denoted by $M \downarrow$. To reconstruct the signal, $x[n]$, all M channels are *upsampled by* M, denoted by $M \uparrow$. Upsampling by M is equivalent to inserting $M-1$ zeros between every sample. This step restores the signal to its original length. Each channel is filtered by reconstruction (or synthesis) filters, and the results are summed to give a reconstructed signal [Meyer 1993]. The process is depicted in Figure 2.6. The analysis and synthesis filters are collectively known as a *filter bank.*

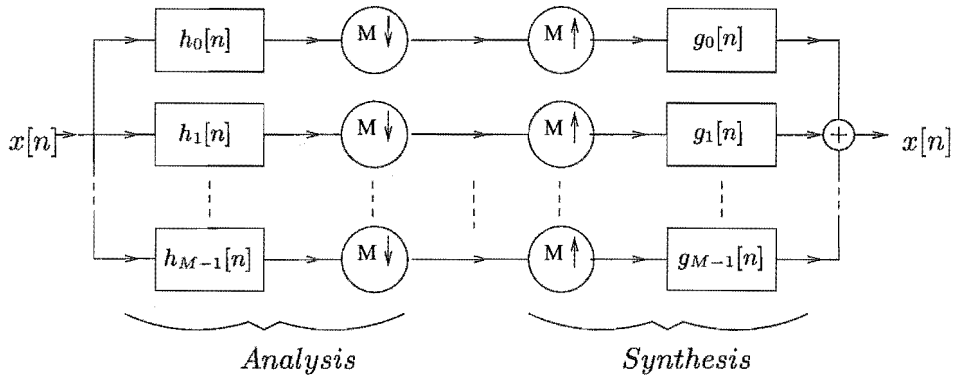


Figure 2.6 M channel analysis and synthesis filter bank.

Consider the case of $M = 2$. The signal is filtered through two frequency channels. Each channel is associated with half the frequency bandwidth of the original. Filtering through two frequency channels is achieved using a low-pass filter, $h_0[n]$, and a high-pass filter, $h_1[n]$. The impulse responses of the filter are given by $h_i[n]$, $i = 0, 1$ and the $h_i[n]$, $n = 0, \ldots, N-1$ are known as *filter coefficients.* The channel with the low-pass filter now contains low frequency information, or, an approximation to the signal, $a[n]$. The channel that filtered the signal with the high-pass filter now contains high frequency information, or, the detail of the signal, $d[n]$. Each channel is now downsampled by 2. If

$h_0[n]$ and $h_1[n]$ are ideal half-band low-pass and high-pass filters respectively, then the original signal will be represented perfectly by the two subsampled versions [Rioul and Vetterli 1991]. Half-band low-pass filtering followed by downsampling by two reduces the resolution by half and doubles the scale [Rioul and Vetterli 1991, Meyer 1993].

Consider the case when the signal, $x[n]$, has been decomposed one step to give $a_1[n]$ and $d_1[n]$. To reconstruct the signal, $\hat{x}[n]$, the two half-band signals, $a_1[n]$ and $d_1[n]$, are upsampled by two and filtered by half-band low-pass, $g_0[n]$, and high-pass, $g_1[n]$, filters respectively, and the filtered signals are then summed. The process is shown in Figure 2.7. The reconstructed signal, $\hat{x}[n]$, is identical to the original signal, $x[n]$, if the filter bank, represented by the filters $h_0[n]$, $h_1[n]$, $g_o[n]$, and $g_1[n]$, is a perfect reconstruction filter bank. Perfect reconstruction is possible when the sub-band analysis and synthesis corresponds to a decomposition onto an orthonormal basis (see Section 2.2.3), followed by a reconstruction given by the summation of the orthogonal projections [Rioul and Vetterli 1991]. This was precisely the case considered in Section 2.2.3.

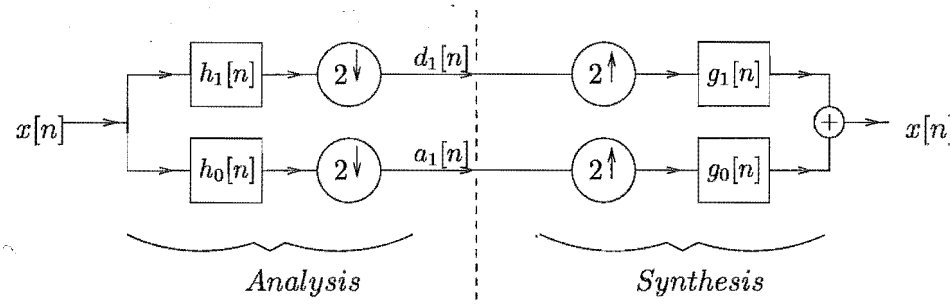


Figure 2.7 A two channel analysis and synthesis filter bank.

2.3.2 Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) is defined in this section. The DWT can be implemented using sub-band coding with a small modification. Sub-band coding traditionally operates at a single level. Wavelet analysis works because the signal is decomposed to multiple scales. To decompose the signal to coarser scales and obtain wider frequency tilings, the decomposition, or analysis, is repeated on the low frequency channel, completing the process. This is illustrated in Figure 2.8. This process leads directly to the construction of Discrete Wavelet Transforms and approximations to Continuous Wavelet Transforms [Mallat 1989, Rioul and Vetterli 1991, Daubechies 1992, Meyer 1993].

The iterated filter bank is used to construct wavelets. Under certain circumstances and after a number of iterations, this discrete system converges to a system where subsequent iterations produce functions that are scaled versions of each other [Rioul and Vetterli 1991, Vetterli and Kovačević 1995].

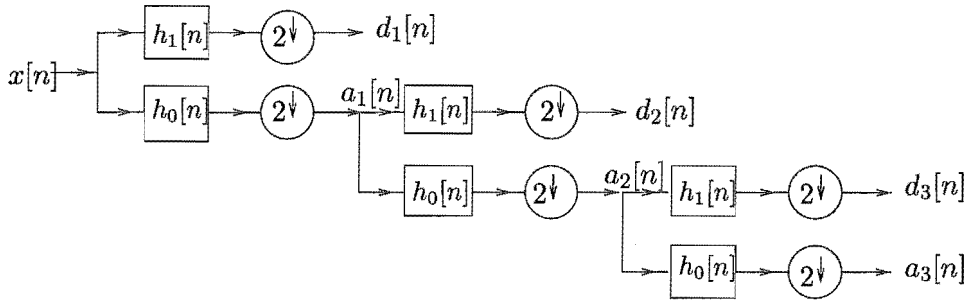


Figure 2.8 A two channel multi-scale analysis and synthesis filter bank.

Two functions are produced when constructing wavelets in the 2-channel case. The first is the function produced by the iteration of the low-pass filter. This function is low-pass, and is the discrete scaling function, ϕ_d. The second function is produced in the same way, except for the last iteration, which is performed using the high-pass filter. This second function is therefore a band-pass function, and is the discrete wavelet function, ψ_d. The scaling function and the wavelet function satisfy the following discrete two-scale difference equations:

$$\begin{aligned} \phi_d[t] &= \sum_{n=-\infty}^{\infty} h_0[n]\phi_d[2t-n], \\ \psi_d[t] &= \sum_{n=-\infty}^{\infty} h_1[n]\phi_d[2t-n]. \end{aligned} \tag{2.29}$$

If the low-pass and high-pass filters form an orthonormal set with respect to even shifts, and the scaling function and wavelet function exist, then the wavelet function forms an orthonormal basis for the set of square integrable functions, $L^2(\mathcal{R})$ [Rioul and Vetterli 1991]. Therefore, the wavelet transform can be used on any finite energy signal.

2.4 SUMMARY

In this chapter, the history and development of wavelets has been described from three perspectives.

The mathematical history of wavelets traces the development of key ideas and concepts that formed the theory of wavelets. The main concepts are: that basis functions exist that, when combined as a weighted sum, can be used to represent any measurable, square integrable n-dimensional function; that these basis functions are continuous; and that these basis functions can be used to analyse the local regularity properties of the function being studied.

Next, the history and the development of wavelets from a signal processing point of view was described. The background is Fourier analysis, with which it is not possi-

ble to analyse the locational information of transient components of a signal from the magnitude of the frequency components. The Short Time Fourier Transform (STFT) is a solution that allows concurrent study of locational and frequency information. However, the accuracy of information from the STFT is limited by the Heisenburg Uncertainty Principle. Wavelet transforms, by using a structured approach to scaling, exploit the Heisenburg Uncertainty Principle. The wavelet transform has parameters of location and scale. The scale parameter is related to frequency. Therefore, analysis of both frequency and location is possible using wavelet transforms. Multiresolutional analysis, which provides the structured approach to scaling, is described. The continuous wavelet transform is defined and its representation in the wavelet domain is explained.

Thirdly, the history and development of wavelets using sub-band coding is described, leading to definitions for the discrete wavelet transform. Thus, while the theory of wavelets is relatively recent, the foundations for the theory have a long history in several fields.

Chapter 3

THE CONSTRUCTION OF WAVELETS

This chapter describes the construction of wavelets using filter banks. The wavelets are used for analysing one-dimensional signals in Chapter 4 and Chapter 5, for developing two-dimensional wavelets in Chapter 6, and analysing ultrasound images in Chapter 8. Filter banks are an integral aspect of constructing wavelets, and a method of representing filter banks, the *polyphase* representation, is described. This representation simplifies the description of the operations of filters in the filter banks. The constraints on these filters are explained, and some examples of wavelets constructed using filter banks are presented at the end of the chapter.

3.1 POLYPHASE REPRESENTATION

This section describes a method of representing signals and the operations of filters on those signals in filter banks. Filter banks were introduced in Section 2.3.1 and examples were shown in Figures 2.5 and 2.6. The method is called *polyphase* representation and is introduced for ease of notation in later sections.

The *polyphase transform* is a method of representing the mathematics of samples of a discrete signal in a system of filter banks. It is often written in the z-domain. The z-transform, $X(z)$, of a discrete function, or sequence, $x[n]$, is given by

$$X(z) = \sum_n x[n]z^{-n}. \tag{3.1}$$

The polyphase transform maps a sequence onto M sequences, with each of the M sequences being a shifted and downsampled version of the original sequence [Vetterli and Kovačević 1995]. The M sequences are represented by a vector of sequences $(x_0[n], x_1[n], \ldots x_{N-1}[n])^T$ where the sequences, or signal polyphase components, are given by

$$x_m[n] = x[Mn + i], \qquad m = 0, ..., M-1. \tag{3.2}$$

Each of the M sequences has a different phase, hence the name polyphase. In the z-domain, the z-transform of the original sequence, $X(z)$, is written as the sum of shifted and upsampled polyphase components:

$$X(z) = \sum_{m=0}^{M-1} z^{-m} X_m(z^M), \tag{3.3}$$

where

$$X_m(z) = \sum_{n=-\infty}^{\infty} x[Mn+m]z^{-n}.$$

For the case $M = 2$, the following five equalities hold:

$$x_0[n] = x[2n], \qquad x_1[n] = x[2n+1],$$

and in the z-domain

$$X_0(z) = \sum_{n=-\infty}^{\infty} x[2n]z^{-n}, \qquad X_1(z) = \sum_{n=-\infty}^{\infty} x[2n+1]z^{-n},$$

and

$$X(z) = X_0\left(z^2\right) + z^{-1}X_1\left(z^2\right). \tag{3.4}$$

The original sequence can be recovered by interleaving the M sequences, and is termed the inverse polyphase transform. In practice this is done by upsampling, inverse shifting, and summing the M sequences [Kovačević and Vetterli 1992, Vetterli and Kovačević 1995, Strang and Nguyen 1996]. The forward and inverse polyphase transform for M phases is shown in Figure 3.1.

It is also necessary to represent the analysis and synthesis filters, $h[n]$ and $g[n]$, in

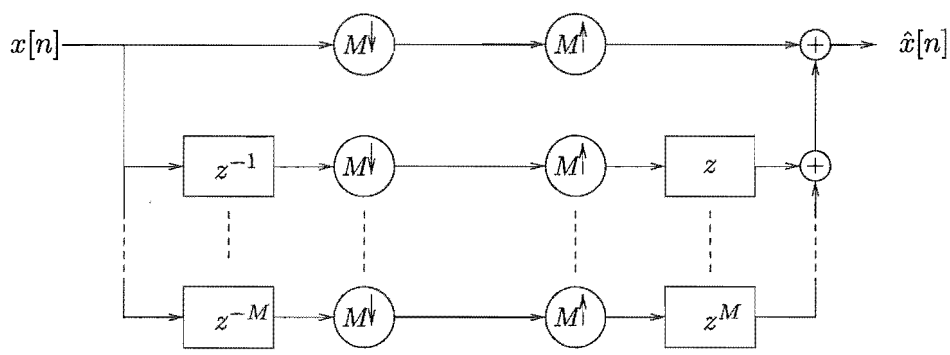


Figure 3.1 Forward and inverse polyphase transform for M phases.

polyphase notation. They are defined to have the inverse phase of the signal, so that when the output of the filter is considered, the output has zero phase, that is, it is the 0^{th} polyphase component of $G(z)H(z)X(z)$. Therefore, $h[n]$ and $g[n]$ are given by the following in the polyphase domain:

$$H(z) = \sum_{m=0}^{M-1} z^m H_m(z^M), \tag{3.5}$$

$$G(z) = \sum_{m=0}^{M-1} z^m G_m(z^M), \tag{3.6}$$

with

$$H_m(z) = \sum_{n=-\infty}^{\infty} h[Mn+m]z^{-n}, \tag{3.7}$$

$$G_m(z) = \sum_{n=-\infty}^{\infty} g[Mn+m]z^{-n}, \quad m = 0, ..., M-1. \tag{3.8}$$

Each filter can also be defined explicitly using its polyphase components:

$$\begin{aligned} H(z) &= \begin{pmatrix} H_0(z) \\ H_1(z) \end{pmatrix} = \begin{pmatrix} H_{00}\left(z^2\right) & H_{01}\left(z^2\right) \\ H_{10}\left(z^2\right) & H_{11}\left(z^2\right) \end{pmatrix} \begin{pmatrix} 1 \\ z^{-1} \end{pmatrix} = H_p\left(z^2\right) \begin{pmatrix} 1 \\ z^{-1} \end{pmatrix}, \\ G(z) &= \begin{pmatrix} G_0(z) & G_1(z) \end{pmatrix} = \begin{pmatrix} 1 & z^{-1} \end{pmatrix} \begin{pmatrix} G_{00}\left(z^2\right) & G_{10}\left(z^2\right) \\ G_{01}\left(z^2\right) & G_{11}\left(z^2\right) \end{pmatrix} = \begin{pmatrix} 1 & z^{-1} \end{pmatrix} G_p\left(z^2\right), \end{aligned} \tag{3.9}$$

where $H_p\left(z^2\right)$ and $G_p\left(z^2\right)$ are known as polyphase analysis and synthesis matrices, respectively, in the upsampled domain (denoted by z^2).

A polyphase representation of a filter bank is shown in Figure 3.2. The output of the filter bank, $\hat{X}(z)$, is given by:

$$\begin{aligned} \hat{X}(z) &= G(z)H(z)X(z), \\ &= T(z)X(z), \end{aligned} \tag{3.10}$$

where $T(z) = G(z)H(z)$.

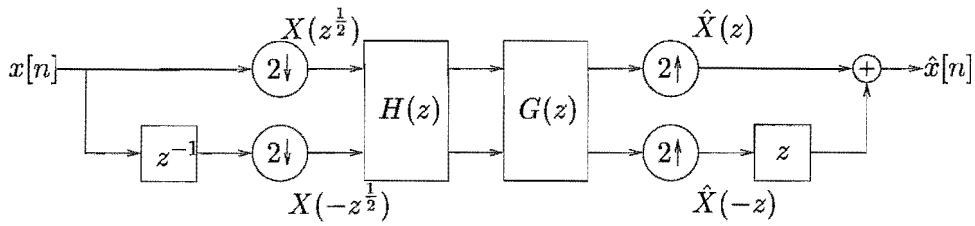


Figure 3.2 Analysis and synthesis filter bank in the polyphase domain for two phases.

The polyphase transform and the subsequent representation of filters and filter banks using polyphase notation have been presented in this section. The method simplifies the description of filter operations on signals.

3.2 PERFECT RECONSTRUCTION FILTER BANKS

In this section, the conditions placed on filters that are required for perfect reconstruction of a signal are described. In addition, filters conditions are described that ensure orthogonality if required.

Perfect reconstruction filter banks occur when the output of the filter bank, $\hat{X}(z)$, is identical to the input of the filter bank, $X(z)$ [Vetterli and LeGall 1989, Vaidyanathan 1990, Vetterli and Herley 1992, Vetterli and Kovačević 1995, Strang and Nguyen 1996]. Perfect reconstruction filter banks are biorthogonal [Strang and Nguyen 1996]. A pure delay of k samples is allowed and occurs when the output of each filter occurs after the corresponding input and is therefore causal [Strang and Nguyen 1996]. Perfect reconstruction is achieved when:

$$T(z) = G(z)H(z) = H(z)G(z) = z^{-k}I, \tag{3.11}$$

where I is the identity matrix.

Inverse matrices automatically involve biorthogonality and are therefore perfect reconstruction systems [Strang and Nguyen 1996]. If $H(z)$ is invertible, then the perfect reconstruction condition may be written as [Vetterli and LeGall 1989, Vaidyanathan 1990, Strang and Nguyen 1996]:

$$G(z) = z^{-k}H^{-1}(z), \tag{3.12}$$

and the filters represented by $H(z)$ and $G(z)$ are biorthogonal.

The filters can be constrained further by considering alias cancellation. Up to this point in this section, only the polyphase form of a quadrature mirror filter has been considered, as in Figure 3.2. To consider alias cancellation, the direct form of a quadrature mirror filter bank is used, as introduced in Section 2.3.1 and shown in Figure 2.7. The direct form given in the z-domain is represented in Figure 3.3. The direct form and polyphase form of a quadrature mirror filter are externally equivalent [Strang and Nguyen 1996], which means that the input and output of the filter banks are equivalent within a pure delay.

In order to produce a filter bank with alias cancellation, a signal and its transformation is followed step-by-step through the direct form of a filter bank. The input signal in the z-domain, $X(z)$ is filtered by the analysis low-pass, $H_0(z)$, and high-pass, $H_1(z)$,

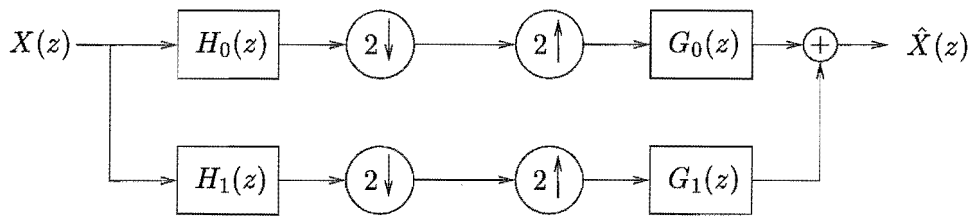


Figure 3.3 Analysis and synthesis filter bank in the direct form for two channels.

filters giving $H_0(z)X(z)$ and $H_1(z)X(z)$ respectively. Each of these is downsampled:

$$\begin{aligned} &\frac{1}{2}\left[H_0\left(z^{\frac{1}{2}}\right)X\left(z^{\frac{1}{2}}\right)+H_0\left(z^{\frac{1}{2}}\right)X\left(z^{\frac{1}{2}}\right)\right] \\ \text{and } &\frac{1}{2}z^{-1}\left[H_1\left(z^{\frac{1}{2}}\right)X\left(z^{\frac{1}{2}}\right)+H_1\left(z^{\frac{1}{2}}\right)X\left(z^{\frac{1}{2}}\right)\right], \end{aligned} \tag{3.13}$$

and then upsampled giving:

$$\begin{aligned} &\frac{1}{2}[H_0(z)X(z)+H_0(-z)X(-z)] \\ \text{and } &\frac{1}{2}z^{-1}[H_1(z)X(z)+H_1(-z)X(-z)]. \end{aligned} \tag{3.14}$$

At this point in the filter bank, aliasing has occurred. Aliasing is manifested as the average of $H_i(z)X(z)$ with $H_i(-z)X(-z)$ [Strang and Nguyen 1996]. In the time domain, the transformed signals have zeros in the odd-numbered components [Strang and Nguyen 1996]. In order to correct for the presence of these zeros, the transformed signal is filtered by the synthesis filters $G_i(z)$, and the transformed signal is now represented in each channel by

$$\begin{aligned} &\frac{1}{2}[G_0(z)H_0(z)X(z)+G_0(z)H_0(-z)X(-z)] \\ \text{and } &\frac{1}{2}[G_1(z)H_1(z)X(z)+G_1(z)H_1(-z)X(-z)]. \end{aligned} \tag{3.15}$$

Finally, the recovered signal $\hat{X}(z)$ is obtained by summing the two channels:

$$\begin{aligned} \hat{X}(z) =&\frac{1}{2}[G_0(z)H_0(z)+G_1(z)H_1(z)]X(z) \\ &+\frac{1}{2}[G_0(z)H_0(-z)+G_1(z)H_1(-z)]X(-z). \end{aligned} \tag{3.16}$$

The aliasing is caused by the term associated with $X(-z)$ [Vaidyanathan 1990, Vetterli and Herley 1992, Strang and Nguyen 1996]. In order for aliasing to be zero, this term is set to zero:

$$\frac{1}{2}[G_0(z)H_0(-z)+G_1(z)H_1(-z)]=0, \tag{3.17}$$

leading to the choices for the synthesis filters of

$$G_0(z) = H_1(-z), \quad \text{and} \quad G_1(z) = -H_0(-z). \tag{3.18}$$

If a pure delay, k, is allowed in the system, and alias cancellation is observed, Equation (3.16) reduces to [Vetterli and LeGall 1989, Vaidyanathan 1990, Strang and Nguyen 1996]:

$$G_0(z)H_0(z) + G_1(z)H_1(z) = 2z^{-k}. \tag{3.19}$$

In the orthogonal case, the high-pass filters are specified by the low-pass filters up to an even shift and a sign change by [Kaiser 1994, Vetterli and Kovačević 1995, Strang and Nguyen 1996]:

$$H_1(z) = -z^{-2K+1}H_0\left(z^{-1}\right), \tag{3.20}$$

leading to the choice of synthesis filters as

$$G(z) = H^T\left(z^{-1}\right). \tag{3.21}$$

Equations (3.19), (3.20) and (3.21) can be expressed in what is known as the *orthonormality condition* [Kaiser 1994]:

$$|H(z)|^2 + |H(-z)|^2 = 2. \tag{3.22}$$

Using the reconstruction condition, Equation (2.26), another condition may be imposed on $H(z)$. This condition is known as the *averaging property* [Kaiser 1994]:

$$H(1) = \sqrt{2}. \tag{3.23}$$

The orthonormality condition and averaging property, Equations (3.22) and (3.23), are two conditions that are used for constructing orthogonal wavelets.

3.3 CONSTRUCTING ORTHOGONAL WAVELETS

This section describes the process of finding filter coefficients for the Daubechies sequence of orthogonal wavelets, ψ^N, where $N = 1, 2, 3, \ldots$, and matching scaling functions, ϕ^N, following Kaiser [1994], Vetterli and Kovačević [1995] and Daubechies [1992].

The filter coefficients associated with a wavelet ψ^N constitute a high pass filter given by $h_1^N[n]$. The filter coefficients associated with the matching scaling function constitute a low pass filter and are given by $h_0^N[n]$. The relationship between the filter coefficients and the wavelet and scaling functions is given in Equation (2.29). The

wavelet and scaling functions are supported on $[0, 2N-1]$. The first member of the sequence, ψ^1 and ϕ^1, is the Haar wavelet and scaling function. As N increases, the wavelets and scaling functions become more regular. For instance, ψ^2 is continuous and ψ^3 is continuously differentiable. However, regularity (described in Section 2.2.4) does not increase linearly with N.

For ease of notation, and to avoid repeating the factor of $\sqrt{2}$, the following notation is used:

$$c[n] = \sqrt{2}h_0[n], \text{ and } d[n] = \sqrt{2}h_1[n],$$

and

$$\begin{aligned} P(z) &= \frac{1}{2}\sum_n c[n]z^{-n} = \frac{1}{\sqrt{2}}h_0(z), \\ \text{and } Q(z) &= \frac{1}{2}\sum_n d[n]z^{-n} = \frac{1}{\sqrt{2}}h_1(z). \end{aligned} \tag{3.24}$$

The strategy is to find polynomials, P(z), that satisfy the orthonormality condition (3.22) and the averaging property (3.23), which can be written as:

$$|P(z)|^2 + |P(-z)|^2 = 1, \quad \text{and } P(1) = 1. \tag{3.25}$$

The *bounding coefficients* are defined as the first and last non-zero coefficients, i.e. $c_0, c_M \neq 0$. All coefficients, c_n, and therefore the wavelet ψ^N and scaling function ϕ^N, are assumed to be real. By the Equations (3.24) and (3.25),

$$\sum_{n=0}^{M} c_n = 2, \quad \sum_{n=0}^{M-2k} c_n c_{n+2k} = 2\delta_k^0, \quad k = 0, 1, \ldots, [M/2], \tag{3.26}$$

where $\delta_k^m = 1$ if $m = k$, and 0 otherwise, and $[M/2]$ is the largest integer less than or equal to $M/2$. M must be odd. To prove this, consider the case of M being even. If M is even, then for $k = M/2$, $c_0 c_M = 0$. This contradicts the assumption that $c_0, c_M \neq 0$. Therefore, M must be odd and the relation $M = 2N - 1$ is taken, with $N \geq 1$, and $[M/2] = N - 1$. Equation (3.26) gives $(N + 1)$ equations in the $2N$ unknowns, $c_0, c_1, \ldots, c_{2N-1}$. When $N = 1$, the unique solution is $c_0 = c_1 = 1$, which is the Haar system. When $N \geq 2$, a further $N - 1$ constraints are added to Equation (3.26). Appropriate constraints are found by using Equation (3.25), which implies that $P(-1) = 0$. If this zero of $P(z)$ at $z = -1$ is assumed to be of order N,

then $P(z)$ can be factored as:

$$P(z) = \left(\frac{1+z}{2}\right)^N W(z) = \left(\frac{1+z}{2}\right)^N \sum_{n=0}^{N-1} w_n z^n, \tag{3.27}$$

where the normalisation has been chosen so that the averaging condition reads $W(1) = 1$. The set of all polynomials satisfying Equation (3.27) can be found using the solution for $N = 1$. The unique solution for $N = 1$ is

$$\begin{aligned} P(z) &= \frac{1+z}{2} = e^{i\pi\omega} cos(\pi\omega), \\ P(-z) &= \frac{1-z}{2} = ie^{i\pi\omega} sin(\pi\omega). \end{aligned} \tag{3.28}$$

To find a solution for any $N \geq 2$, the abbreviations

$$\begin{aligned} C(z) \equiv cos(\pi\omega) &= \frac{z^{*\frac{1}{2}} + z^{\frac{1}{2}}}{2}, \\ \text{and } S(z) \equiv sin(\pi\omega) &= \frac{z^{*\frac{1}{2}} - z^{\frac{1}{2}}}{2i} \end{aligned} \tag{3.29}$$

are used. The identity $C^2 + S^2 = 1$ is raised to the power $2N - 1$:

$$(C^2 + S^2)^{2N-1} = \sum_{k=0}^{2N-1} \frac{M!}{k!(M-k)!} C^{4N-2-2k} S^{2k}. \tag{3.30}$$

The sum on the right-hand side contains $2N$ terms. Let $\mathcal{P}_N(C)$ be the sum of the first N terms. Using the identity $S^2 = 1 - C^2$, the polynomial of degree $4N - 2$ is defined:

$$\mathcal{P}_N(C) \equiv \sum_{k=0}^{N-1} \frac{M!}{k!(M-k)!} C^{4N-2-2k} (1 - C^2)^k \geq 0. \tag{3.31}$$

If a polynomial $P(z)$ can be found such that $|P(z)|^2 = \mathcal{P}_N(C)$, then $|P(-z)|^2$ is automatically the sum of the last N terms from Equation (3.30) and the orthonormality condition in Equation (3.25) is satisfied because $C(-z) = -S(z)$ and $S(-z) = C(z)$.

Equation (3.31) can be factored in the form

$$\mathcal{P}_N(C) = C^{2N} \mathcal{W}_N(S) = \left|\frac{1+z}{2}\right|^{2N} \mathcal{W}_N(S), \tag{3.32}$$

where

$$\mathcal{W}_N(S) = \sum_{k=0}^{N-1} \frac{M!}{k!(M-k)!} (1 - S2)^{N-1-k} S^{2k} \geq 0, \tag{3.33}$$

and $\mathcal{W}_N(0) = 1$. In order to find $P(z)$ in the form of Equation (3.27), it is necessary to find a square root, $W(z)$, of $\mathcal{W}_N(S)$, that is,

$$W(z) = \sum_{n=0}^{N-1} w_n z^n \text{ such that } |W(z)| = \sqrt{\mathcal{W}_N(S)}, \tag{3.34}$$

with $W(1) = 1$ and real coefficients w_n. For the case $N = 2$, the Equation (3.33) is

$$\mathcal{W}_2(S) = 1 + 2S^2 = 2 - \frac{z^* + z}{2}. \tag{3.35}$$

$W(z) = a + bz$ is a first order polynomial with $a, b \in \mathcal{R}$. Therefore, using Equation (3.34), the solution is

$$a^2 + b^2 = 2,\ 2ab = -1 \Rightarrow (a+b)^2 = 1,\ (a-b)^2 = 3, \tag{3.36}$$

which leads to the Daubechies filters with $N = 2$:

$$\begin{aligned} P_{\pm}(z) &= \frac{1}{2}\left(\frac{1+z}{2}\right)^2 \left[(1 \pm \sqrt{3}) + (1 \mp \sqrt{3})z\right], \\ &= \frac{1}{8}[(1+\sqrt{3})z^3 + (3+\sqrt{3})z^2 + (3-\sqrt{3})z + 1 - \sqrt{3}]. \end{aligned} \tag{3.37}$$

This filter is the 4-tap Daubechies' filter, daub2, within a phase shift and a scale factor of $1/\sqrt{2}$. The phase shift ensures causality. The scale factor normalises the output.

Daubechies filters for $N > 2$ can be constructed using the same method. Solving for $W(z)$ becomes more difficult as N increases.

3.4 CONSTRUCTING LINEAR PHASE WAVELETS

Linear phase wavelets result in less phase distortion than orthogonal wavelets. Our visual system is more sensitive to phase distortions than amplitude distortions, so when applying wavelet based techniques to signals and images for visual analysis, it is usually preferable to use linear phase wavelets [Daubechies 1992, Kovačević and Vetterli 1992, Vetterli and Herley 1992, Vetterli and Kovačević 1995, Misiti *et al.* 1996].

Wavelets can be described by their symmetry. Non-symmetrical wavelets are not symmetrical about their mid-point. Symmetrical wavelets have coefficients of which one half are the mirror image of the other half. Anti-symmetrical wavelets have coefficients of which one half are the negative mirror image of the other half [Daubechies 1992, Kaiser 1994, Vetterli and Kovačević 1995, Strang and Nguyen 1996]. Anti-symmetry is also referred to as odd symmetry. Linear phase wavelets are symmetrical, or at least

anti-symmetrical. Hence, filters of linear phase wavelets may be written as:

$$H(z) = \pm z^{-K+1} H\left(z^{-1}\right). \tag{3.38}$$

Linear phase wavelets are constructed so that they are biorthogonal. Therefore, the analysis and synthesis filter pairs form dual wavelet bases and the corresponding filter banks have perfect reconstruction [Daubechies 1992].

3.4.1 Sizes of Linear Phase Wavelets

The length of a filter is given by the number of samples in the impulse response of the filter. The sizes of linear phase filters dictate their symmetry [Vaidyanathan 1990, Vetterli and Herley 1992, Kaiser 1994, Vetterli and Kovačević 1995]. The following cases hold for linear phase perfect reconstruction real FIR filters $H_0(z)$ and $H_1(z)$:

1. Both filters are symmetric and of odd lengths where the lengths differ by an odd multiple of 2.
2. One filter is symmetric and the other anti-symmetric. Both lengths are even, and are equal or differ by an even multiple of 2.
3. One filter is of odd length, and the other is of even length. Both filters have all their zeros on the unit circle. Either they are both symmetric, or one is symmetric and the other is anti-symmetric.

In order to prove these statements, take a polynomial $P(z) = H_0(z)H_1(-z)$. Since $H_0(z)$ and $H_1(-z)$ are both linear phase, $P(z)$ is also linear phase. $P(z)$ may have either an odd or an even length.

If $P(z)$ is an odd length, then the filters are either both an odd length, or both an even length. Taking the difference between $P(z)$ and its negative, $P(-z)$, gives an expression for the coefficients, p_N,:

$$\begin{aligned} P(z) - P(-z) &= \sum_{l=0}^{2N} p_l z^{-l} \left[1 - (-1)^l\right], \\ &= \sum_{l=0}^{N-1} 2p_{(2l+1)} z^{-(2l+1)}. \end{aligned} \tag{3.39}$$

Since $P(z)$ is symmetric, both $P(z)$ and $P(-z)$ are symmetric or anti-symmetric about the point $l = N$. Therefore, $P(z) - P(-z)$ is symmetric about this point and the only non-zero coefficient is the central one, p_N. Therefore, $P(z)$ is symmetric which implies that $H_0(z)$ and $H_1(-z)$ are both either symmetric or anti-symmetric. In the following, the length of a filter is given by L_i. For the two possible cases of the filters being either both of odd length or both of even length, the two cases below hold:

1. L_0 and L_1 are both odd. Therefore, (L_0+L_1-1) is odd. The center of symmetry, which is $(L_0+L_1)/2-1$ samples from the end points, must be an odd-numbered sample. Therefore, $(L_0+L_1)/2 = L_0+(L_0-L_1)/2$ must be even. Hence, $(L_0-L_1)/2$ is odd, and the length difference (L_0-L_1) is an odd multiple of 2. $H_0(z)$ and $H_1(-z)$ cannot both be anti-symmetric because the two polyphase components are also anti-symmetric when the length is odd and perfect reconstruction is not possible [Vetterli and Herley 1992]. There are no solutions of the same length, therefore, if $H_0(z)$ and $H_1(z)$ are both symmetric and have odd lengths, their lengths must differ by an odd multiple of two.

2. L_0 and L_1 are both even. $(L_0+L_1)/2-1$ must be odd so the center term is not in the same set as the end terms. $L_0+(L_0-L_1)/2$ must be even and since L_0 is even, the difference between the lengths must be an even multiple of two. When the length of $H_i(z)$ is even, then $H_i(-z)$ has opposite symmetry. Therefore, the even length solution leads to a symmetric/anti-symmetric pair.

If $P(z)$ is even, then one filter has even length and the other has odd length. Now,

$$\begin{aligned} P(z)-P(-z) &= \sum_{l=0}^{2N-1} p_l z^{-i}\left[1-(-1)^l\right], \\ &= \sum_{l=0}^{N-1} 2p_{(2l+1)} z^{-(2l+1)}. \end{aligned} \tag{3.40}$$

$P(z)$ has only a single non-zero odd-indexed coefficient, but odd-indexed coefficients are paired with even-indexed coefficients. Hence $P(z)$ has only one even-indexed coefficient of the form

$$P(z) = p_l z^{-l}\left(1 \pm z^{2N-1-2l}\right). \tag{3.41}$$

Since $P(z)$ has all its zeros on the unit circle (at the $2(N-j)-1$ roots of ± 1), $H_0(z)$ and $H_1(-z)$ also have all zeros on the unit circle. If $H_0(z)$ and $H_1(z)$ are both anti-symmetrical, both must have zeros at $z=1$. This forces $H_1(-z)$ to have a zero at $z=1$ or $z=-1$ depending on whether the filter is odd or even length, respectively. This implies that $P(z)$ has a double zero at $z=1$ or a pair at $z=1$ and $z=-1$. Since $P(z)$ contains only the $2(N-l)-1$ roots of ± 1, both possibilities are ruled out. Therefore any filter pair, one of odd length and the other of even length, must have opposite symmetry.

3.4.2 Linear Phase and Orthogonal Wavelets

One of the properties of orthogonal wavelets is that, in general, they distort phase information. It is well known that, for the FIR two-channel case, there are no real,

perfect reconstruction filters that are both orthogonal and linear phase, except the Haar filters [Smith and Thomas P. Barnwell 1987, Vetterli and LeGall 1989, Vaidyanathan 1990, Vetterli and Herley 1992, Vetterli and Kovačević 1995, Strang and Nguyen 1996]. This can be shown by the following, which is taken from Vetterli and Kovačević [1995].

Using Equation (3.21) with zero delay, orthonormality implies

$$H(z)H^T\left(z^{-1}\right) = I, \tag{3.42}$$

which in turn implies

$$H_{00}(z)H_{00}\left(z^{-1}\right) + H_{01}(z)H_{01}\left(z^{-1}\right) = 1. \tag{3.43}$$

The filters are of an even length, from Equation (3.26). From the previous section, a pair of even length linear phase filters have one that is symmetrical and one that is anti-symmetrical. Take the low-pass filter to be symmetrical and using the condition for linear phase, Equation (3.38), then

$$\begin{aligned} H_0(z) &= H_{00}\left(z^2\right) + z^{-1}H_{01}\left(z^2\right), \\ &= z^{-K+1}H_0\left(z^{-1}\right), \\ &= z^{-K+1}\left[H_{00}\left(z^{-2}\right) + zH_{01}\left(z^{-2}\right)\right], \\ &= z^{-K+2}H_{01}\left(z^{-2}\right) + z^{-1}\left[z^{-K+2}H_{00}\left(z^{-2}\right)\right]. \end{aligned} \tag{3.44}$$

Therefore, the polyphase components are related as:

$$H_{00}(z) = z^{-K/2+1}H_{01}\left(z^{-1}\right), \quad H_{01}(z) = z^{-K/2+1}H_{00}\left(z^{-1}\right). \tag{3.45}$$

Substituting the second part of Equation (3.45) into Equation (3.43) results in:

$$H_{00}(z)H_{00}\left(z^{-1}\right) = \frac{1}{2}. \tag{3.46}$$

The only FIR, real polynomial satisfying Equation (3.46) is:

$$H_{00}(z) = \frac{1}{\sqrt{2}}z^{-l}. \tag{3.47}$$

Similarly, $H_{01} = \frac{1}{\sqrt{2}}z^{-k}$, and the low-pass and high-pass filters are given by

$$H_0(z) = \frac{1}{\sqrt{2}}\left(z^{-2l} + z^{-2k-1}\right), \quad H_1(z) = H_0(-z), \tag{3.48}$$

and the only solution ($l = k = 0$) yields the Haar filters, or trivial variations of the Haar filters.

3.4.3 Odd Length Linear Phase wavelets

The construction of odd length linear filters is discussed in this section. A pair of odd length filters is chosen with both filters symmetrical.

It is possible to use the properties of symmetric polynomials to create a cascade specifically for linear phase filters. One property used is that two symmetric polynomials multiplied together result in a third symmetric polynomial [Kovačević and Vetterli 1992]. Also used is the property that two symmetric polynomials of the same size summed together result in a third symmetric polynomial [Kovačević and Vetterli 1992]. Symmetric polynomials may be written $z^{-K+1}p\left(z^{-1}\right) = p(z)$. A cascade can be written in the form

$$H_p''(z) = H_p(z)H_p'(z), \tag{3.49}$$

where

$$H_p(z) = \begin{pmatrix} a(z) & b(z) \\ c(z) & d(z) \end{pmatrix}, \tag{3.50}$$

and $a(z), b(z), c(z)$, and $d(z)$ are all symmetric polynomials. For pairs of filters, each of odd length, but differing in length by two, the following restrictions are also placed on $H_p(z)$; $Det[H_p(z)] = z^{-K}$ with $degree[a(z)] = degree[d(z)] = K$, $degree[b(z)] = K-1$, and $degree[c(z)] = K+1$. For $K = 1$, which produces symmetric filters of lengths 3 and 5, the following example holds:

$$H_p(z) = \begin{pmatrix} 1+z^{-1} & a \\ 1+bz^{-1}+z^{-2} & c(1+z^{-1}) \end{pmatrix}, \tag{3.51}$$

where, for perfect reconstruction, the determinant must equal a pure delay, that is, $Det[H_p(z)] = z^{-1}$. Therefore choose $a = c$, with $a \neq 0$, and $b \neq 2$.

In the general case, using Equations (3.49) and (3.51), for $2K+1$ and $2K+3$ length linear phase filters, the following result holds [Vetterli and LeGall 1989]:

$$H_p(z) = \prod_{k=1}^{K} \begin{pmatrix} 1+z^{-1} & a_k \\ 1+b_k z^{-1}+z^{-2} & a_k(1+z^{-1}) \end{pmatrix}, \quad a_k \neq 0, b_k \neq 2. \tag{3.52}$$

It has been shown that the smallest solutions, with $K = 1$, for both the orthogonal and linear phase filters are complete in one-dimension [Vetterli and LeGall 1989]. It has also been shown that the smallest size perfect reconstruction filter pairs are complete in any number of dimensions [Kovačević and Vetterli 1992].

3.5 CONSTRUCTING WAVELETS FROM FILTER COEFFICIENTS

Wavelets and their scaling functions are constructed from their filter coefficients using Equation (2.29). The process is illustrated in Figure 3.4. Constructing a scaling function requires three steps:

1. determine low pass reconstruction filter coefficients;
2. upsample by 2; and
3. convolve with the original low pass reconstruction filter.

This process produces the scaling function, $\phi_d[n]$, at one scale. The scaling function at the next scale is obtained by iterating the process. Iteration is performed by passing the result from the previous scale back into step two. In this way, the scaling function can be obtained for any desired scale.

Wavelets are constructed in a similar manner by:

1. determining the high pass reconstruction filter coefficients;
2. upsampling by 2; and
3. convolving with the original low pass reconstruction filter.

The wavelet, $\psi_d[n]$, is obtained at the desired scale by repeating steps two and three, always passing the result from the previous iteration back into step two.

Wavelets and their scaling functions converge to a continuous function. The order of continuity is defined by the regularity of the wavelet [Vetterli and Herley 1992, Misiti *et al.* 1996], as given in Section 2.2.4.

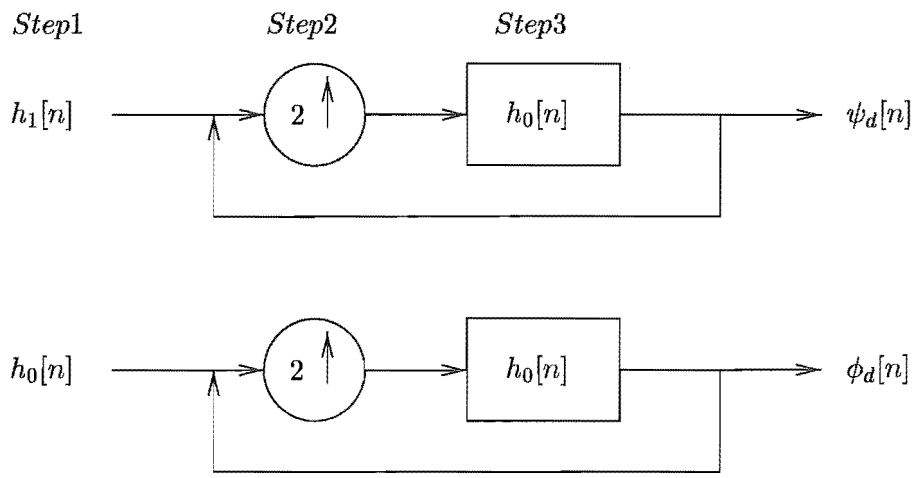


Figure 3.4 Filter bank representation for building wavelets, $\psi_d[n]$, and scaling functions, $\phi_d[n]$. See the text for a description of the steps.

3.6 EXAMPLES OF WAVELETS AND THEIR SCALING FUNCTIONS

Filter coefficients associated with the Daubechies wavelet families, denoted daubN, biorthogonal wavelet families, denoted biorN1.N2, Coifman wavelet family, denoted coifN, and the symlet wavelet family, denoted symN, are found in the literature [Daubechies 1992, Kaiser 1994, Misiti *et al.* 1996, Strang and Nguyen 1996]. N denotes the order of the wavelet. The daubN, biorN1.N2, coifN and symN wavelets will be compared in the next section for their suitability for analysing infant breathing signals. Some examples of filter coefficients and their resulting wavelet and scaling functions are given in this section.

The impulse response of the low pass filter for the Daubechies orthogonal wavelet daub2 is given in Table 3.1. In the orthogonal case, the high pass filter is found by reversing the low pass filter, then multiplying every even coefficient by (-1). The reconstruction filters are found by reversing the coefficients. In the z-domain, the filters are related to the low pass filter, $H_0(z)$, using Equations (3.18), and (3.20), by:

$$\begin{aligned} H_1(z) &= -H_0\left(-z^{-1}\right), \\ G_0(z) &= H_0\left(z^{-1}\right), \\ G_1(z) &= H_1\left(z^{-1}\right) = -H_0(-z). \end{aligned} \tag{3.53}$$

In the time domain the filters are related to the low pass filter, $h_0[n]$, of length N, by:

$$\begin{aligned} h_1[n] &= (-1)^n h_0[N-n-1], \\ g_0[n] &= h_0[N-n-1], \\ g_1[n] &= h_1[N-n-1] = (-1)^n h_0[n]. \end{aligned} \tag{3.54}$$

Filter Coefficient	Orthogonal Daubechies daub2
$h_0[0]$	0.3415
$h_0[1]$	0.5915
$h_0[2]$	0.1585
$h_0[3]$	-0.0915

Table 3.1 Impulse responses for daub2 low pass analysis filter.

The results for the Daubechies scaling function and wavelet, daub2, for the first six iterations of the algorithms given in Section 3.5 are shown in Figure 3.5. Note the quick convergence to a self-similar function, with the scaling and wavelet functions in

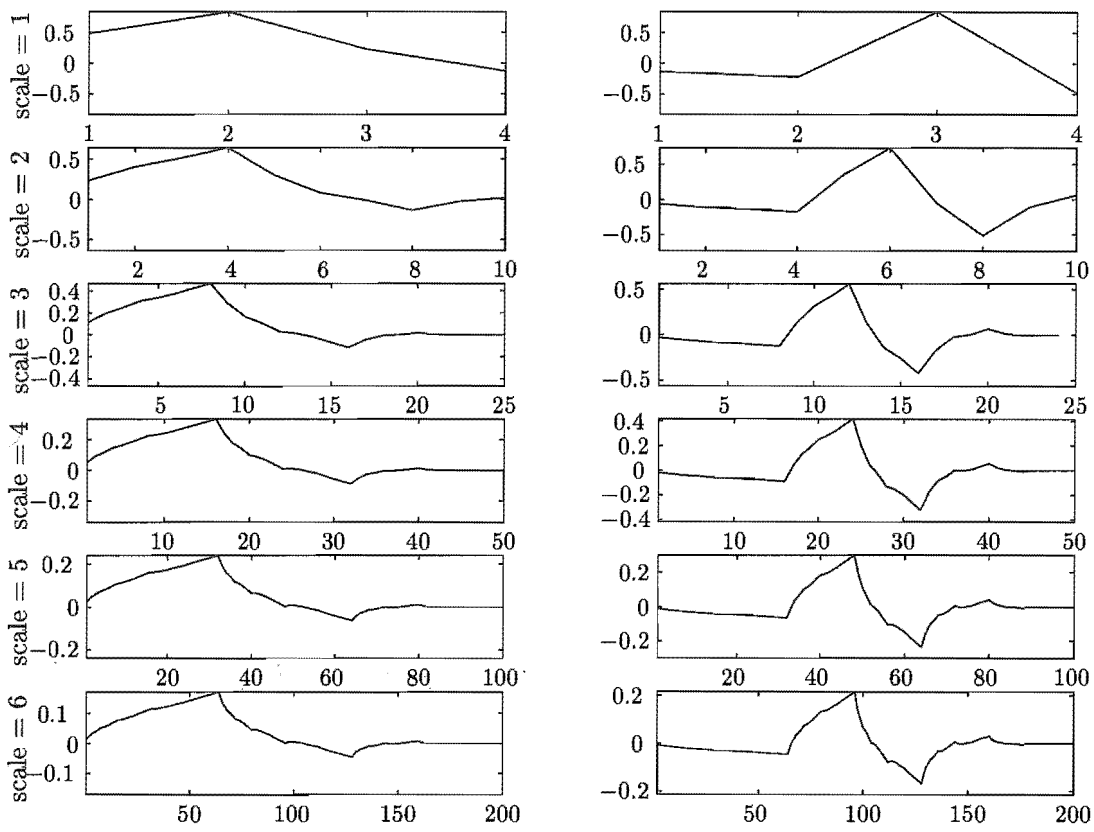


Figure 3.5 The first six iterations of the daub2 scaling function (left) and wavelet (right).

the last two iterations being almost identical. The daub2 wavelet is continuous, but none of its derivatives are continuous. Therefore, it has a low regularity.

The impulse responses of the dual wavelet and scaling function pair for the Bior6.8 filters are given in Table 3.2. Each pair has odd length filters differing in length by an odd multiple of two. Therefore, as expected from Section 3.4.1, all the filters are symmetrical. All four filters are given.

The results for the Biorthogonal scaling functions and wavelets, bior6.8, after the first six iterations of the algorithms given in Section 3.5 are shown in Figure 3.6. Note the symmetry and the smoothness of the filters, indicating a higher degree of regularity than the daub2 wavelet.

3.7 SUMMARY

In this chapter, aspects of constructing wavelets and scaling functions are explained with some examples of individual wavelets and scaling function pairs given.

The polyphase representation of signals and filter banks is described. This representation is used to detail the constraints necessary for perfect reconstruction filter banks. The analysis and synthesis filters are defined in relation to each other so that

n	$h_0[n]$	$h_1[n]$	$g_0[n]$	$g_1[n]$
0	0	0	0	0
1	0.0019	0	0	-0.0019
2	-0.0019	0	0	-0.0019
3	-0.0170	0.0144	0.0144	0.0170
4	0.0119	-0.0145	0.0145	0.0119
5	0.0497	-0.0787	-0.0787	-0.0497
6	-0.0773	0.0404	-0.0404	-0.0773
7	-0.0941	0.4178	0.4178	0.0941
8	0.4208	-0.7589	0.7589	0.4208
9	0.8259	0.4178	0.4178	-0.8259
10	0.4208	0.0404	-0.0404	0.4208
11	-0.0941	-0.0787	-0.0787	0.0941
12	-0.0773	-0.0145	0.0145	-0.0773
13	0.0497	0.0144	0.0144	-0.0497
14	0.0119	0	0	0.0119
15	-0.0170	0	0	0.0170
16	-0.0019	0	0	-0.0019
17	0.0019	0	0	-0.0019

Table 3.2 Impulse responses for bior6.8 low pass analysis, $h_0[n]$, high pass analysis, $h1[n]$, low pass synthesis, $g_0[n]$, and high pass synthesis, $g_1[n]$, filters.

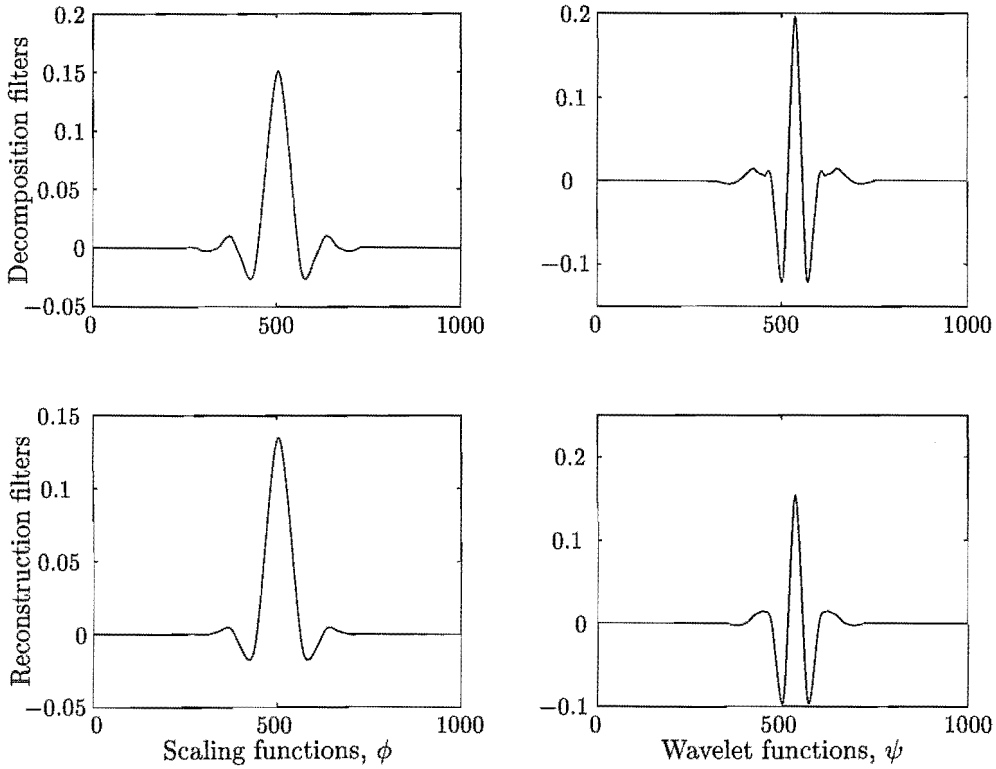


Figure 3.6 The sixth iteration of the bior6.8 decomposition scaling function (top-left) and wavelet (top-right) and reconstruction scaling function (bottom-left) and wavelet (bottom-right).

alias cancellation is achieved. Orthogonal filter banks are additionally constrained such that the low-pass and high-pass filters are specified by each other up to an even shift and a sign change.

Construction of one-dimensional orthogonal and linear phase wavelets is described with specific design examples given. Constraints on the symmetry of linear phase wavelets are explained in relation to the size of the filters.

The techniques presented here allow for the development of several different wavelets that may be used in a variety of applications. As well as giving an understanding of the nature of wavelets, some of the wavelets are constructed and used for analysis in later chapters. The following chapter compares the effectiveness of different wavelets for analysing infant breathing signals.

Chapter 4

SELECTING A WAVELET TO ANALYSE AMPLITUDE MODULATION OF INFANT BREATHING SIGNALS

This chapter describes the process of selecting a wavelet to analyse infant breathing signals with particular emphasis on assessing amplitude modulation. The characteristics of infant breathing signals are described in the first section. In the second section, test signals that are simplified representations of breathing signals are developed, and wavelet transforms of these signals are used to illustrate how to interpret signals in the wavelet domain. The third section discusses the traditional approaches to selecting a wavelet and some of the drawbacks with these approaches. The fourth section develops a new method for selecting a wavelet to analyse infant breathing signals. The fifth section presents the results of the selection process. Conclusions are drawn in the final section of this chapter.

4.1 CHARACTERISTICS OF INFANT BREATHING SIGNALS

Infant breathing signals are often studied in the context of Sudden Infant Death Syndrome (SIDS). In this section, the characteristics of breathing signals are described as they relate to breathing patterns that have been associated with SIDS [Haidmayer *et al.* 1982a]. Specifically, quiet sleep breathing patterns have been linked to fatal events.

Three breathing patterns have been linked to SIDS. The first is apnoea, which is prolonged pause in breathing. The second pattern is periodic breathing, which is a succession of apnoea. The third pattern, which can be considered a generalization of periodic breathing, is amplitude modulation [Preiss *et al.* 1975, Rantonen *et al.* 1998]. In periodic breathing, the amplitude modulation is so severe that breathing ceases altogether, resulting in a succession of apnoea. Sometimes, as well as these three breathing patterns associated with SIDS, there are large inspiratory efforts followed by short pauses in breathing, which are associated with sighs. However, even though these patterns may occur during quiet sleep, the majority of breathing during quiet sleep is

of constant amplitude (relative to other sleep states) and with only small variations in frequency. The patterns associated with SIDS are the focus of this study. Specifically, the characteristics of breathing signals during quiet sleep that are pertinent are:

1. one main frequency with only small variations relating to the primary breathing frequency,
2. amplitude modulation of various magnitudes at one or more frequencies that are slower than the primary breathing frequency,
3. short, high frequency components relating to large inspiratory efforts and followed by pauses associated with sighs.

The aim of studying breathing of infants at high risk for SIDS and normal infants is to provide measurable differences of the various characteristics of breathing between the two groups. Overall characteristics of the breathing including the median, mean and variability of infant breathing rates and breath to breath intervals have been successfully measured and studied [Richards *et al.* 1984, Schechtman *et al.* 1988, Schechtman *et al.* 1996]. However, these measures are not able to assess the impact of periodic breathing or amplitude modulation on a signal. The number and duration of periodic breathing components have been evaluated from polysomnographic studies [Kelly *et al.* 1980, Richards *et al.* 1984, Richard *et al.* 1998], but instances of smaller amplitude modulation were not measured. Investigations of periodic breathing in infants have included spectral analysis [Jahnukainen *et al.* 1995], with one publication [Rantonen *et al.* 1998] analysing low frequency (0.03 - 0.17 Hz) and high frequency (0.3 - 1.3 Hz) components of the respiratory activity. The low frequency components of the spectrum do give an indication of the amplitude modulation of the signal. However, for precise measurement of amplitude modulation, spectral analysis based on the Fourier Transform is problematic, as there is an assumption that a stationary signal is being analysed. Since the pertinent breathing characteristic, specifically amplitude modulation, is non-stationary, a time-frequency analysis is appropriate to accurately quantify the effect on the primary breathing frequency component of the signal.

Wavelet analysis is a technique of time-frequency analysis that has been successfully applied to analysis of other physiological signals [Unser and Aldroubi 1996], including breathing signals [Akay *et al.* 1996]. The application of the continuous wavelet transform to non-stationary signals allows each component of the signal to be identified by frequency and locality. Once the individual components of a signal have been identified, it is possible to assess the impact of each component on the signal.

In the case of infant breathing signals during quiet sleep, the chief area of interest is the extent to which the amplitude modulation of the primary breathing frequency influences the breathing, as periodic breathing is associated with SIDS. The more general pattern of amplitude modulation has also been linked to SIDS [Rantonen *et al.* 1998],

however, this characteristic has not been widely studied. The remainder of this chapter examines the features of breathing characteristics in the wavelet domain. The choice of wavelet to analyse the breathing is discussed. The next section gives examples of these characteristics and describes the corresponding features in the continuous wavelet domain.

4.2 APPLICATION OF ONE-DIMENSIONAL WAVELETS TO TEST SIGNALS

This section describes the application of wavelet transforms to test signals that highlight pertinent characteristics of breathing signals, as presented in Section 4.1. The characteristics include short but large inspiratory efforts and periodic features of varying frequency. Based on these characteristics, delta functions and sinusoids of varying frequencies have been chosen as test signals. Each test signal has one characteristic, or a carefully combined set of characteristics, so that the reader may examine how the individual characteristics in the time domain are represented in the wavelet domain, and also how a combination of characteristics in the time domain interact in the frequency domain. The intention is to make it easier for the reader to identify the individual characteristics in the wavelet transforms of breathing data.

The selection of a wavelet to analyse amplitude modulation of infant breathing signals is discussed in Section 4.3. The analyses of the test signals are presented using one of the wavelets assessed in Section 4.3 to be an appropriate wavelet for analysing breathing signals.

The examples are illustrated in figures with the original signal above and the corresponding transform below, as in Figure 4.1. The scale of the graph representing each test signal is set to encompass the minimum and maximum of the signal. The test signals are all 256 samples long, and the nature of the signals is described in the text and in the figure titles.

The continuous wavelet transforms were obtained using the Matlab Wavelet Toolbox, version 1.2. The results are shown with the lowest scales at the bottom of the plot. The lowest scales correspond to the highest frequencies of the signal. The scales that were calculated for each transform are indicated in the title of the transform in the form "a = [*first scale : step between scales : final scale*]." The intensity of each plot is calculated using the absolute value of the wavelet transform coefficients. Using the absolute value ensures the lowest intensity values (black) occur at zero crossings, making it simple to gauge the frequency of a signal for a given scale in the wavelet transform. The intensity is either normalised over all the coefficients, or normalised for each scale. Normalising the intensity over all the coefficients allows the relative power of the coefficients to be compared between scales. Normalising the intensity for each scale allows the frequency information to be highlighted for every scale. The method

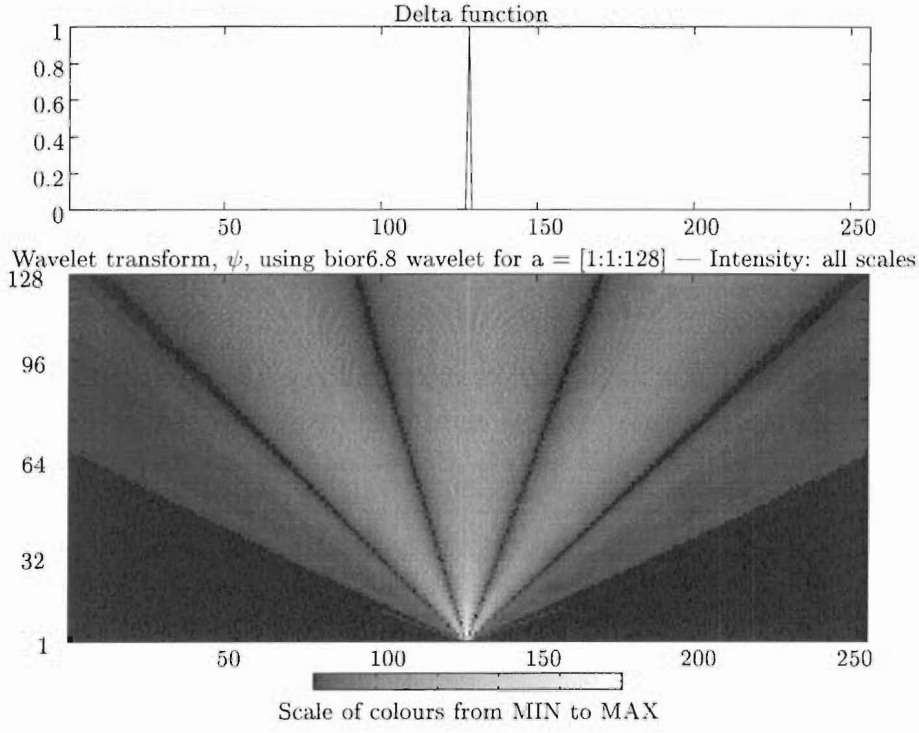


Figure 4.1 Example 1: Delta function and associated continuous wavelet transform with intensity normalised over all coefficients.

of normalisation used for each figure is indicated in the title for the transform as "all scales" for normalisation over all the coefficients, or "by scales" for normalisation over each scale.

Example 1: Delta Function

A delta function and its continuous wavelet transform is shown in Figure 4.1. The intensity is normalised over all the wavelet transform coefficients. There is good locational information at low scales: the delta function is located to within 10 samples in the first five scales. The first 20 scales are also of higher intensity than the rest of the transform coefficients. This indicates that the energy of the signal is concentrated at lower scales, and therefore higher frequencies, as expected for a delta function. There is spreading of the information across all samples at scales above 70. Therefore, there is no information regarding the location of the delta function at higher scales. The fanning effect is due to the lengthening window size of the wavelet transform as the scales increase.

Figure 4.2 presents a different display of the same analysis of the delta function. The intensity has been normalised for each scale as opposed to over all the wavelet transform coefficients in Figure 4.1. Compared to Figure 4.1, the pattern created by the delta function is enhanced, with a higher intensity over more scales. The frequency information is more ap-

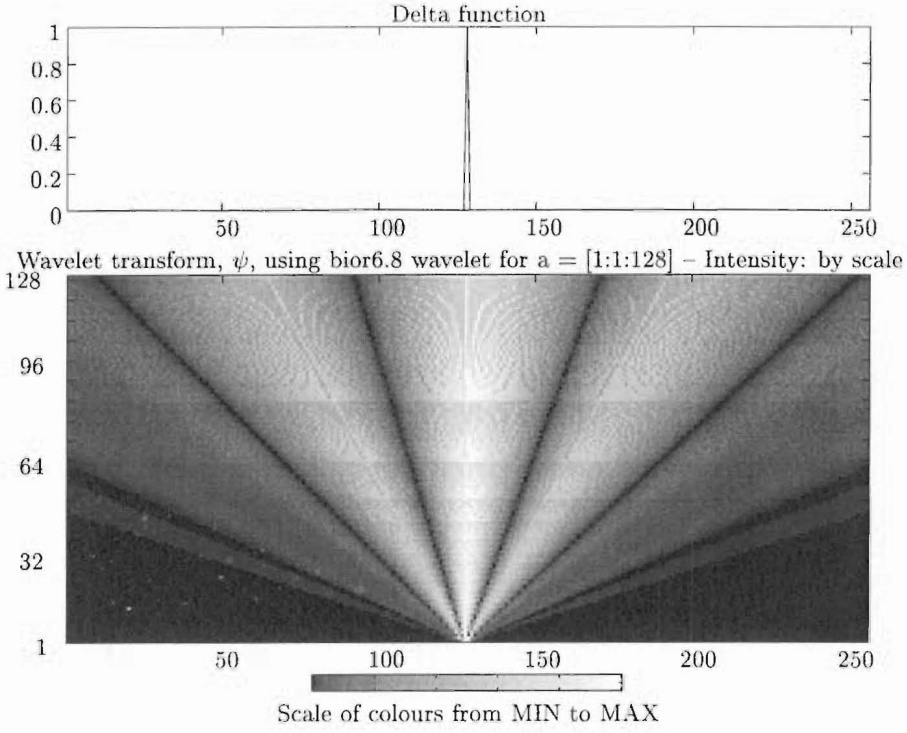


Figure 4.2 Example 1: Delta function and associated continuous wavelet transform with the intensity normalised for each scale.

parent using normalisation for each scale, however, information regarding the relative importance of the transform at different scales is lost. When information regarding the relative importance of the frequency components is desired, the intensity is normalised over all the coefficients.

Example 2: Sinusoid

The second test signal is a single sinusoid, which is shown with its transform in the wavelet domain in Figure 4.3. The intensity of the transform is normalised over all scales. The transform contains dark vertical strips that represent the zero crossings of the sinusoid. The bright vertical columns represent the peaks and troughs of the original signal. The strips and columns occur between scales nine to 45, with the brightest part of the columns occurring between the scales of 22 and 32. The brightest part of the columns corresponds to the scales at which the wavelet basis function most closely resembles the test signal.

When analysing signals, frequency information is often required, as described in Section 1.1. By taking a horizontal line through the transform and measuring from dark strip to dark strip, the period of the transformed sinusoid can be obtained. If the absolute value of the coefficients is used in the transform intensity calculation, the period must be doubled. This is because a sinusoid has two zero crossings per cycle, and hence, there

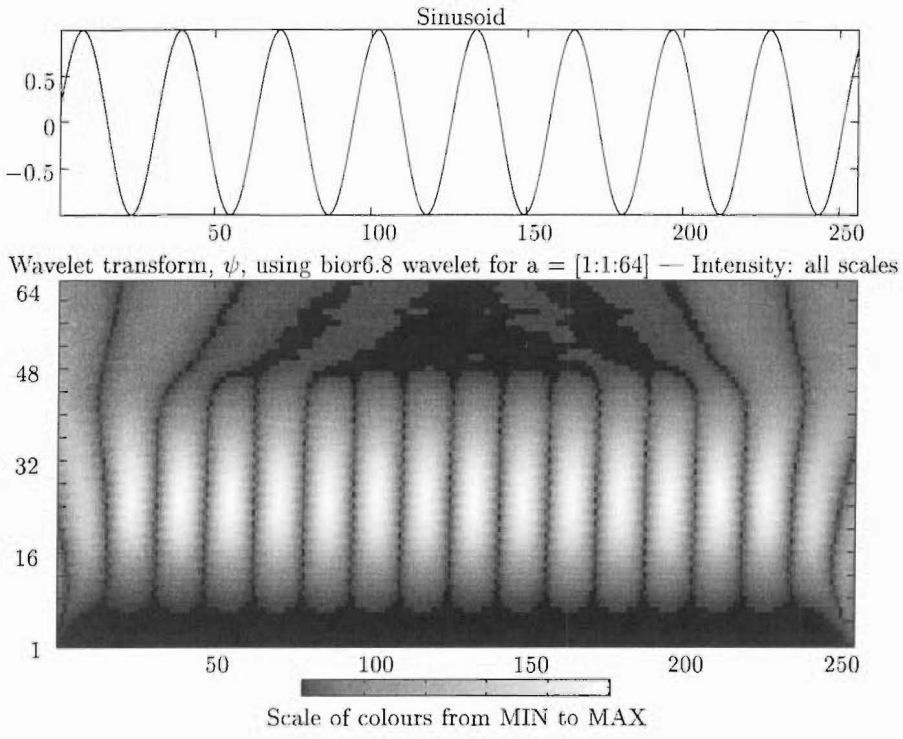


Figure 4.3 Example 2: Sinusoid function and associated continuous wavelet transform.

are two dark strips per cycle. Therefore, to obtain the period of the signal, the period between strips is measured and doubled. The frequency is calculated by taking the inverse of the period. The ideal scale at which to measure the period is where the columns are brightest, because this is where the frequency component is represented best in the transform.

Edge effects can be seen in Figure 4.3, appearing as a fan effect from each edge. The fan effect due to an impulse was seen in Figure 4.1. The edge effects cause the pattern of the sinusoid to be disturbed at the top of the wavelet transform, above scales of 45 and for approximately 75 samples in from each edge. The edge effects are also apparent as curving of the two vertical dark strips closest to each edge. These edge effects are distortions of the true wavelet transform, which is calculated over an infinite number of samples – in the same way that side lobes are present in spectra calculated using the Fourier Transform on finite length signals.

Example 3: Sinusoid with delta function

Figure 4.4 displays a sinusoid combined with a delta function, and the wavelet transform of this combined signal. The intensity of the transform is normalised over all coefficients. Both the position of the delta function and the frequency of the sinusoid can be seen, with a small fan at the bottom of the transform and vertical strips seen in the mid-range of

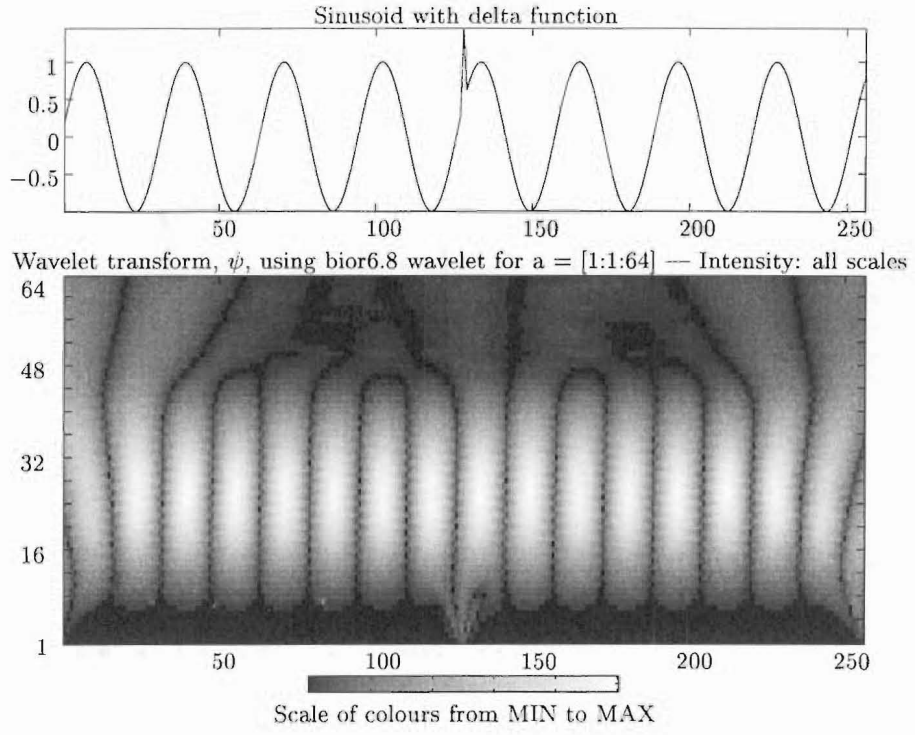


Figure 4.4 Example 3: Sinusoid function with added delta function and associated continuous wavelet transform.

scales. The location of the delta function is distinct for scales up to five, slightly less than for the delta function on its own. The delta function and sinusoid function interact in the transform between the scales of five and 13 and between the scales of 45 and 49. The interaction appears as a blending of the two characteristics with the ends of the sinusoid pattern and the fan effect modifying each other. Even though the delta function does influence all scales as seen in Figure 4.1, the effect of the delta function is masked by the effect of the sinusoid at the middle range of scales between 13 and 45—the scales at which the sinusoid pattern was observed in Figure 4.3. This indicates that different frequencies dominate different scales, and hence, that different frequencies can be isolated by analysing the appropriate scales. The sinusoid information in the transform is the same as for Figure 4.3. As expected, the intensity of the transform indicating the presence of the delta function is reduced, because the signal's energy is split between the sinusoid and the delta function. The edge effects are the same as in Example 2, affecting the outer two dark strips at each edge and the top of the sinusoidal pattern above scales of 45.

Example 4: Three adjacent sinusoids

This example consists of three sinusoids of different frequencies that are adjacent but not overlapping in time. The signal and its transform are

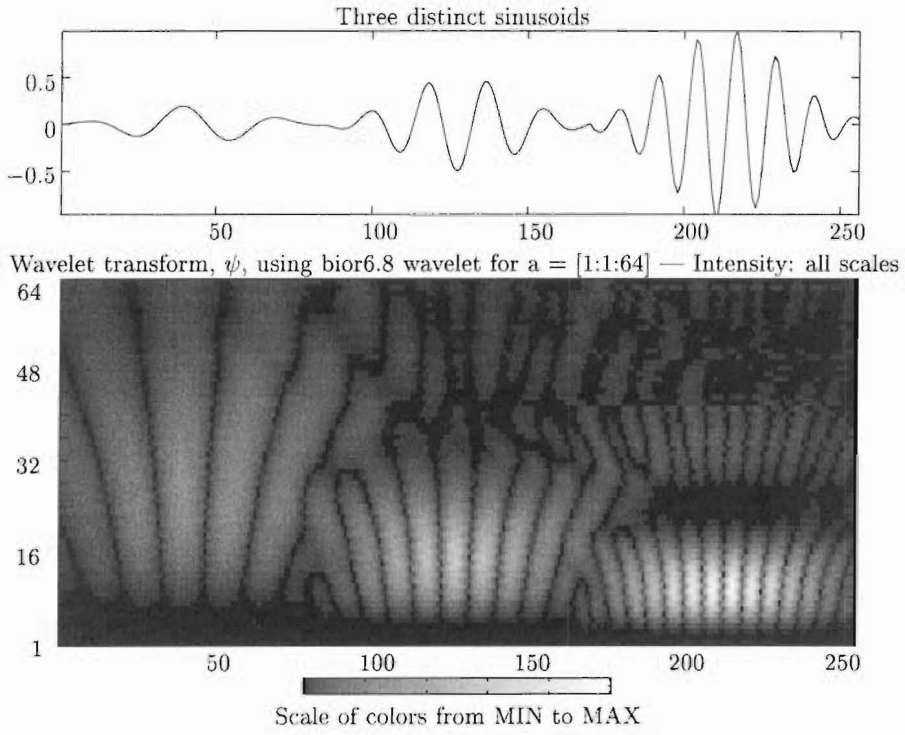


Figure 4.5 Example 4: Three Sinusoids with frequencies $\frac{1}{2\pi}$, $\frac{3}{2\pi}$, and $\frac{5}{2\pi}$ separated in time and windowed by a Hamming Window to reduce discontinuities at the junctions, and associated continuous wavelet transform.

shown in Figure 4.5. The first sinusoid in the signal has a frequency of $\frac{1}{2\pi}$, located in the first 85 samples. The second sinusoid has a frequency of $\frac{3}{2\pi}$, located in the samples 86 to 190. The third sinusoid has a frequency of $\frac{5}{2\pi}$, located in the samples 191 to 256. Each section of the signal is windowed by a Hamming Window to reduce discontinuities at the junctions between frequency components.

The separate frequencies can be identified in the transform. The first sinusoid occupies approximately the first 85 samples and scales nine to 64. The middle sinusoid occupies approximately the middle 85 samples and scales six to 37. The last sinusoid occupies approximately the last 85 samples and scales four to 21. The higher frequencies occupy a band of lower scales, whilst the lower frequencies occupy a band of higher scales. Each section of the signal exhibits a fan-shaped effect as one frequency stops and the next starts. There are interference effects due to harmonics of the frequency components. The harmonic effects are evident above the bands that show the frequency components lower in the transform.

Example 5: Three concurrent sinusoids

This example consists of three sinusoids summed together across all 256 samples. The signal and its transform are shown in Figure 4.6. Both

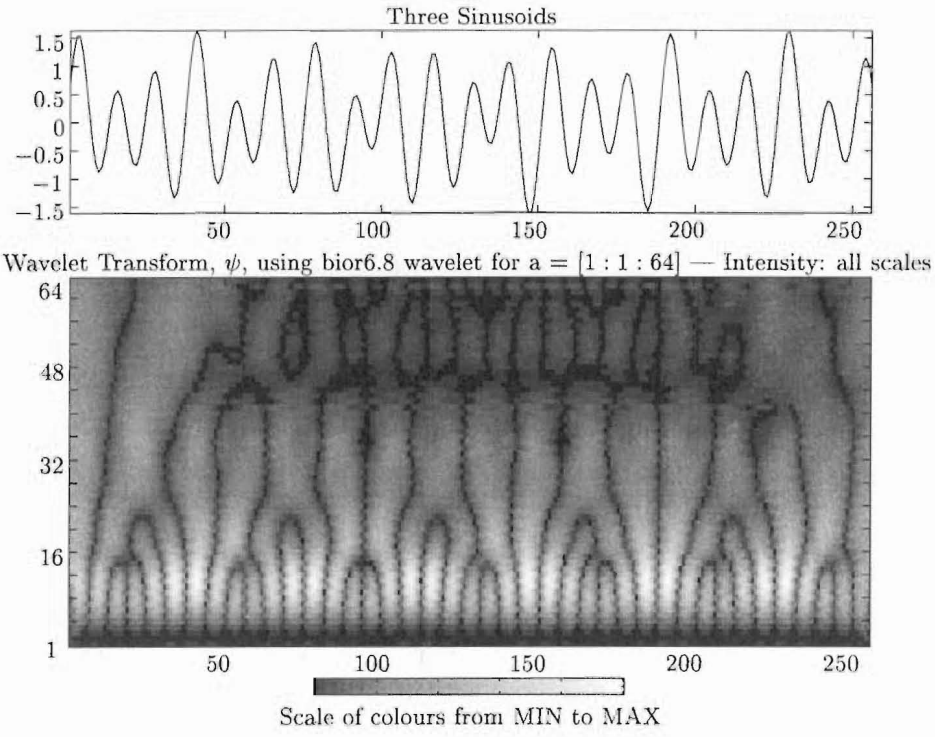


Figure 4.6 Example 5: Three Sinusoids with frequencies $\frac{1}{2\pi}$, $\frac{3}{2\pi}$, and $\frac{5}{2\pi}$, and associated continuous wavelet transform.

the test signal and its transform in this example are starting to resemble breathing signals and their transforms. The three sinusoids in the test signal have the same frequencies as the sinusoids in Example 4. The signal is also the same as the signal in Figure 2.1 (c) with its Fourier Transform in Figure 2.1 (a). The separate frequencies can be identified in the continuous wavelet transform in bands between the scales of five and 13 for the high frequency component, between the scales of 17 and 22 for the mid frequency component and between the scales of 25 and 41 for the low frequency component. These ranges are narrower than the positionally separate frequency components in Example 4, but they do still allow separation of the frequency components. The transform has less intensity at higher scales than the single sinusoid in Example 2, due to the energy being concentrated in the higher frequency components of the signal. There are also interference effects in scales above scale 41. The interference effects at lower scales, evident in Example 4, are masked by the transform due to the lowest frequency component.

Summarising, the five test signals are simplified representations of characteristics encountered in infant breathing signals. Consequently, the resulting wavelet transforms are simplified examples of wavelet transforms of breathing data. Delta functions tend to exhibit fan-shaped transforms, and pure sinusoids exhibit transforms with vertical

bands of light and dark. Different frequencies within the same signal can be isolated by selecting appropriate scales. Breathing signals, such as those analysed in Chapter 5, can be considered to comprise of combinations of sinusoids and delta functions. The test examples demonstrate that when a combination of different frequency components is present in the time domain, the visibility of individual characteristics in the wavelet domain is reduced, but still present. Typically, each characteristic is apparent over a narrower range of scales. The test examples are used as a reference when interpreting patterns seen in wavelet transforms of breathing signals.

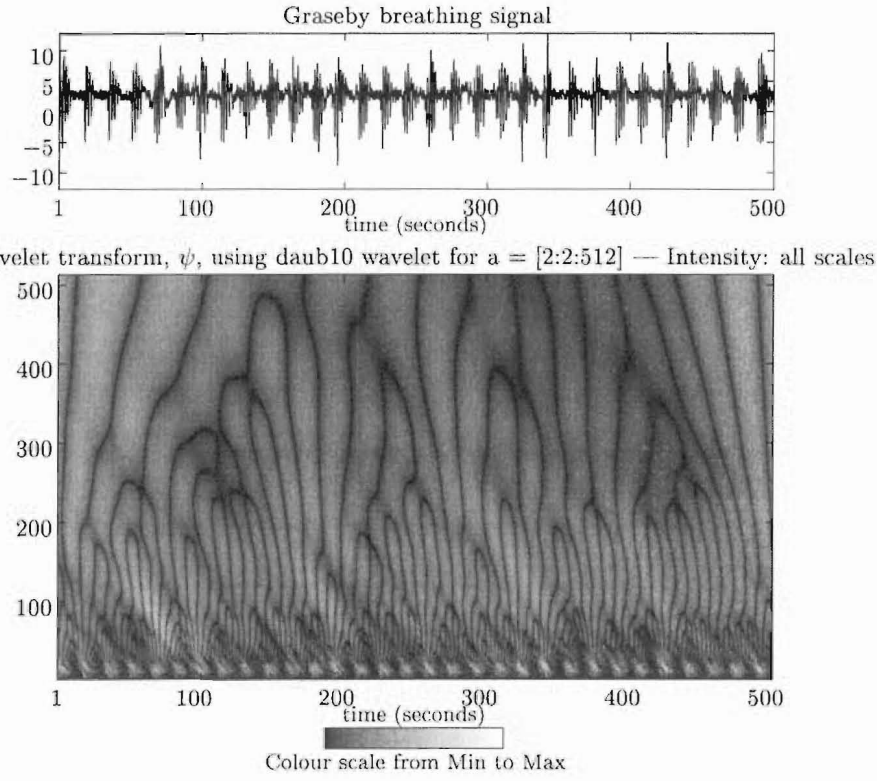
Having created a reference set of examples, it is necessary to choose a suitable wavelet to analyse the breathing signals. The effect of the choice of wavelet is considered as, for a given signal, every wavelet has a different response in the wavelet domain.

4.3 SELECTION OF A WAVELET FOR ANALYSING BREATHING SIGNALS

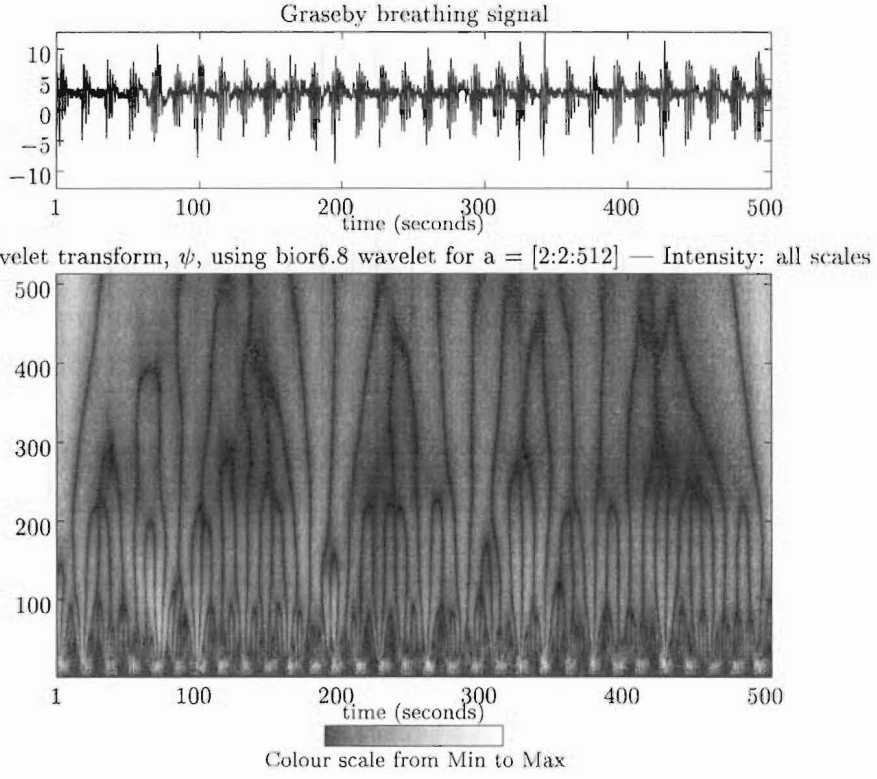
There are a variety of wavelets, each of which have different characteristics, and therefore, selecting a wavelet should be made with the particular analysis in mind. Recalling from Section 4.1, the pertinent characteristics of infant breathing signals are: 1) one main frequency with only small variations relating to the primary breathing frequency; 2) amplitude modulation of various magnitudes at one or more frequencies that are slower than the primary breathing frequency; and 3) short, high frequency components relating to large inspiratory efforts and followed by pauses associated with sighs.

There are three main considerations when analyzing these signals: 1) coincidence across scales to maintain temporal alignment; 2) time-frequency localization across a broad range of frequencies; and 3) regularity characteristics, due to the range of frequencies over which the signal will be analysed. Traditionally, these requirements are modelled by the following methods. Coincidence across scales is maintained by using symmetric, or linear phase, wavelets [Meyer 1993]. Bounds for time-frequency localisation can be specified using the uncertainty principle [Rioul and Vetterli 1991, Unser *et al.* 1992]. The number of vanishing moments of a filter is directly related to the filter regularity [Rioul 1993]. These requirements are discussed in more detail below.

The first consideration is coincidence across scales. Linear phase wavelets, or near linear phase wavelets, minimise the distortion of positional information [Daubechies 1992, Vetterli and Herley 1992, Sun and Sclabassi 1995, Vetterli and Kovačević 1995, Misiti *et al.* 1996] and ensure coincidence across scales. This choice is illustrated by applying the orthogonal daub10 wavelet and the linear phase bior6.8 wavelet to a breathing signal, as in Figure 4.7. The two wavelets have similar support widths of 20 and 17, respectively. The breathing signal is a section of an infant's breathing signal with 5000 samples, representing 500 seconds of breathing. The dark stripes in the daub10 wavelet



(a) The continuous wavelet transform resulting from an orthogonal wavelet, daub10, applied to a breathing signal.



(b) The continuous wavelet transform resulting from a linear phase wavelet, bior6.8, applied to a breathing signal.

Figure 4.7 Comparing the result of orthogonal and linear phase wavelets applied to a breathing signal. Both wavelets have a similar support width.

transform are not vertical, but skewed horizontally as the scales increase. The skewing in the transform as a result of using the daub10 wavelet is pronounced. The linear phase wavelet produces a cleaner and more easily interpreted transform. The transform is easier to interpret because the lines indicating the zero crossings are straight, which means the locational information is preserved. Hence, for analysing modulation of breathing signals in the wavelet domain, a linear phase wavelet or near linear phase wavelet should be used.

The second consideration is time–frequency localisation. Time–frequency localisation is governed by the Heisenburg Uncertainty Principle, as outlined in Section 2.2.2. The uncertainty principle bound is given by, $\Delta t \Delta f$ (Equation (2.3)), where $(\Delta t)^2$ is the variance of the power distribution of the wavelet and $(\Delta f)^2$ is the variance of the power spectrum of the Fourier transform of the wavelet [Bracewell 1986]. The uncertainty principle bound corresponds to the area of the tiles in the time–frequency plane (Figure 2.3). The wavelet with the smallest tile area, or uncertainty principle bound, will have the best time-frequency localisation. This criterion has been used to select wavelets for biomedical applications [Unser and Aldroubi 1996].

The third consideration is regularity, described in Section 2.2.4. Functions with more widely varying characteristics can be analysed with a wavelet that has greater regularity [Drouiche and Kateb 1999], and therefore, regularity can be used as a selection criterion for wavelets [Mojsilovic *et al.* 1998]. The number of vanishing moments of a wavelet is related to the regularity [Mojsilovic *et al.* 1998], so a high number of vanishing moments suggests a better wavelet. The signals being considered here will be analysed to identify the location of significant frequency components and not to identify discontinuities or other higher order features. Therefore, the regularity of the wavelet is less important for this application. However, in order to match the behaviour of the signals being analysed, continuous wavelets are used because of the continuous nature of breathing signals.

These three considerations can be quantified for various wavelets, but they do not completely specify the behavior of any given wavelet when applied to different types of signals. These methods rely only on the characteristics of the wavelets themselves, and not on the characteristics of the interaction of the wavelets with the signal being analysed. Therefore, a new set of tests is proposed that measures characteristics of the transform of the wavelets and the signal.

As seen in Section 4.2, a wavelet exhibits different features in the wavelet domain, depending on the characteristics of the signal under study. It is also interesting to note that different wavelets applied to the same signal can result in different parameters of the features in the wavelet domain. Therefore, to test the different wavelets, a test signal is selected that isolates certain characteristics of the breathing signals to be studied.

As stated in Section 4.1, the characteristics of breathing during quiet sleep that are pertinent to the studies in this thesis are:

1. one main frequency with only small variations relating to the primary breathing frequency,

2. amplitude modulation of various magnitudes at one or more frequencies that are slower that the primary breathing frequency,

3. short, high frequency components relating to large inspiratory efforts.

In order to provide measurable differences of the various characteristics of breathing, three steps are performed. The first is calculating the frequency components of the signal; the second is determining the location and duration of the frequency components; and the third is calculating the relative importance of each frequency component.

Since there is more than one frequency that needs to be isolated, good localisation in the frequency domain is required. Good localisation in the temporal domain is also required to provide accurate measurements of both the timing and duration of frequency components.

To test the performance of different wavelets in response to a breathing signal, test signals similar to those used to illustrate the behaviour of various features in the wavelet domain in Section 4.2 are used. Using these test signals, new performance tests are developed and used to compare the results between wavelets.

4.4 WAVELET PERFORMANCE MEASURE

Four tests are developed that each assess a different aspect of the wavelet performance in the wavelet domain [Macey and Page 2000]. The first two tests primarily assess frequency localisation. The second two tests primarily assess temporal localisation, the ability of the wavelet to accurately locate a frequency component in time.

The differences between wavelet responses are quantified using the energy of the wavelet transform. The total energy, E_ψ, of the wavelet transform, $X(t,j)$, is given by summing the energy for each scale, j, and time, t, of the transform:

$$E_\psi = \frac{\sum_{j=1}^{J}\sum_{t=1}^{N}|X(t,j)|^2}{\|x\|_2^2}. \tag{4.1}$$

where $\|x\|_2^2$ is the norm of the signal. The discussion is limited to those wavelets that are *scale-limited* [Odegard *et al.* 1992]. The notion of a scale-limited wavelet is that the total energy of the wavelet transform of a given signal will be acceptably small for

scales $j > J$, represented by:

$$E_{\psi} = \frac{\sum_{j=J}^{\infty} \sum_{t=1}^{N} |X(t,j)|^2}{\|x\|_2^2} \leq \epsilon, \tag{4.2}$$

where ϵ is a tolerance. Therefore, the energy in the wavelet domain of a wavelet that is scale limited for a given signal may be considered finite.

From Equation 4.1, scalar energy, the normalized energy of the wavelet transform calculated across time at each scale, $E_{\psi,t}(j)$, at scale j is given by:

$$E_{\psi,t}(j) = \frac{\sum_{t=1}^{N} |X(t,j)|^2}{\|x\|_2^2}. \tag{4.3}$$

Similarly, temporal energy, the normalized energy of the wavelet transform calculated across scales at each time, $E_{\psi,j}(t)$, at time t is given by:

$$E_{\psi,j}(t) = \frac{\sum_{j=1}^{J} |X(t,j)|^2}{\|x\|_2^2}. \tag{4.4}$$

Different aspects of the energy in the wavelet domain are used to assess the performance of wavelets. Test signals are constructed to isolate features in the wavelet domain that are to be optimised.

4.4.1 Rate of Decay of Scale Harmonics

For a scale-limited wavelet, given a single frequency signal, the ideal response in the wavelet domain is a single band of energy covering a narrow range of scales. The band of energy relates to the signal frequency. In non-ideal circumstances, *scale harmonics* occur. The definition of scale harmonics is analogous to the definition of harmonics in the frequency domain. Scale harmonics are described using the scalar energy (see Equation (4.3)). The *fundamental scale*, j_f, is the scale at which the first (and main) peak relating to a particular frequency occurs in scalar energy. Scale harmonics occur at odd multiples of the fundamental scale $(2n+1)j_f$, where n is the n^{th} harmonic and $n \in \mathcal{Z}$, where $\mathcal{Z}$ is the set of integers. Scale harmonics decay towards zero for a scale-limited signal. An example is shown in Figure 4.8.

For the purposes of analysis, it is desirable for the scale harmonics to rapidly reduce to an acceptably low percentage of the fundamental peak height. This ensures that any interference with other frequencies in the wavelet domain will be minimal. Hence, a performance measure based on the rate of decay of the scale harmonics is constructed.

The rate of decay of scale harmonics is calculated in the following manner. Initially, the peaks are identified at the fundamental and scale harmonics. The peak heights are then expressed as a ratio of the peak at the fundamental scale, or fundamental peak

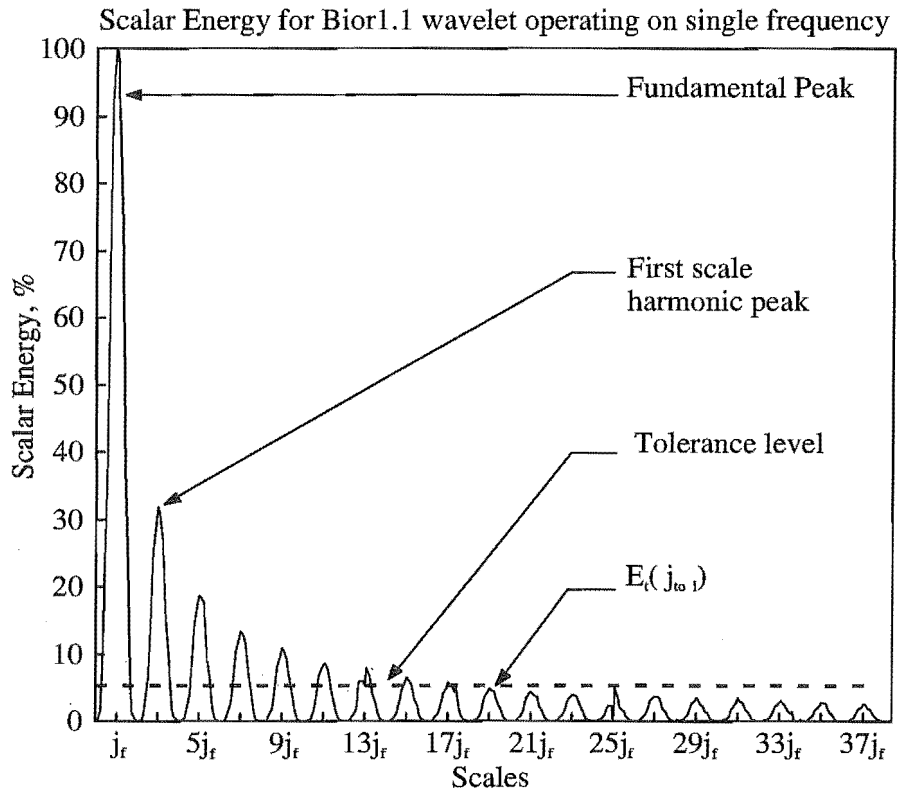


Figure 4.8 Normalised scalar energy of the wavelet transform of a single frequency signal ($N = 6000, f = 0.5$ Hz, sampling rate = 10 Hz) using the Biorl.1 wavelet. Salient features are the fundamental scale, j_f, the subsequent scale harmonics at $(2n+1)j_f$, the (representative) tolerance level and the peak at which the peak heights drop below the tolerance level, $E_t(j_{tol})$.

height, $E_t(j_f)$. For a fast rate of decay, the difference between the fundamental height and the height of the first peak under a prescribed tolerance should be maximised. The first peak height under the tolerance, $E_t(j_{tol})$, should occur at the lowest harmonic possible. The difference between the fundamental peak height and the height of the first peak under the tolerance, expressed as a ratio of the fundamental peak height, is divided by the harmonic number of the tolerance peak, n_{tol}, to produce a ranking, $R_{harmonic}$, based on the scale harmonic behaviour of the wavelet:

$$R_{harmonic} = \frac{E_t(j_f) - E_t(j_{tol})}{n_{tol} E_t(j_f)} \tag{4.5}$$

$R_{harmonic}$ approaches 1 as the tolerance peak height reduces to zero and the tolerance harmonic number approaches 1. The performance increases as $R_{harmonic}$ approaches 1.

4.4.2 Slope of Scalar Energy

The second criterion ranks wavelets on the frequency localisation of their response to a single frequency signal. The frequency localisation is assessed using the steepness of

the energy drop-off from the fundamental peak. The faster the energy drops off, the smaller the range that the peak covers in the scale dimension, and consequently, the finer the frequency resolution in the wavelet domain. To compare the performance of different wavelets, scalar energy is normalised relative to the maximum scalar energy, i.e.,

$$E_t(j)_{norm} = \frac{E_t(j)}{\max\left(E_t(j)\right)}. \tag{4.6}$$

The maximum slope of each side of the normalized fundamental scalar energy peak is used to rank the wavelets, R_{slope}:

$$R_{slope} = \max\left[\left|\frac{dE_t(j)_{norm}}{dj}\right|\right]\Big|_{j=1}^{j=2j_f}. \tag{4.7}$$

Ideally, the slope would approach infinity, relating to the frequency being perfectly isolated. However, due to the Heisenburg Uncertainty principle and the discretisation of scale, the maximum value for R_{slope} is 1.

4.4.3 Cumulative Temporal Energy

The third criterion is developed to measure the symmetry of each wavelet response to a delta function, as well as the localisation of each response. Cumulative energy, $E_C(t)$, is calculated from the time of the impulse to the edges of the signal, individually on the left and right sides of the impulse. Therefore:

$$E_{C_{left}}(t) = \sum_{t_{delta}}^{t=1} \frac{\sum_{j=1}^{J} |X(t,j)|^2}{\|x\|_2^2} \tag{4.8}$$

and

$$E_{C_{right}}(t) = \sum_{t_{delta}}^{N} \frac{\sum_{j=1}^{J} |X(t,j)|^2}{\|x\|_2^2} \tag{4.9}$$

The cumulative energy for each time from the impulse is evaluated as a percentage of the total cumulative energy. The cumulative energy response is assessed at a given time, t_{CE}, from the impulse by the percentage of energy reached at that time. The time, t_{CE}, is chosen to be close to the impulse response since the ideal response has good temporal resolution. Ideally, these percentages are identical on each side and as large as possible. This ensures a symmetrical response as well as one that is constrained in time, thereby having good temporal resolution. The process for ranking wavelets using the cumulative energy is, firstly, to take the difference, D, between the percentages on the left and right sides, p_l and p_r, respectively. Secondly, the percentages are combined,

and finally, since the difference should be small and the combined percentage, C, should be large, the cumulative energy ranking, R_{CE}, is given by subtracting the difference from the combined percentage:

$$\begin{aligned} D &= |\, p_l - p_r \,|\,, \\ C &= p_l + p_r, \\ R_{CE} &= C - D. \end{aligned} \tag{4.10}$$

As the wavelet response approaches the ideal case, R_{CE} approaches 100%.

4.4.4 Uniformity Criteria

The fourth assessment criterion is based on the response of the wavelets to a signal composed of two frequencies. One frequency occurs only during the first half of the signal and a second frequency only during the second half. The transition is continuous, but the derivatives are not.

The assessment of the wavelet transforms of the dual frequency signal is based on uniformity criteria. The method involves measuring the time that any frequency component is present in the wavelet domain over a defined range of scales. This is accomplished by isolating the scale ranges that show significant frequency components based on scalar energy. The temporal energy is calculated separately for each of those ranges. Uniformity criteria are then applied to each range limited temporal energy profile.

The uniformity criteria are based on the height and frequency of peaks within the temporal energy. Frequency is considered to be uniform when at least four consecutive peaks have similar spacing. Uniform frequency sections with an average height greater than a tolerance are selected. The start and end times of these sections give the approximate time periods that the frequency component is present, as illustrated in Figure 4.9. The duration of the sections, T_U, is expressed as a percentage of the actual duration of the frequency components in the signal. The fourth assessment criterion, $R_{uniformity}$, is the average of the percentage duration for each frequency component:

$$R_{uniformity} = \frac{\sum_{n=1}^{N} T_U(n)}{N}, \tag{4.11}$$

where N is the number of frequency components in the signal. For the test signal used, the ideal case is 100%.

4.4.5 Overall Wavelet Performance Measure

Each of the above measures quantifies an aspect of performance of a wavelet, with higher values corresponding to better performance. To assess the overall performance

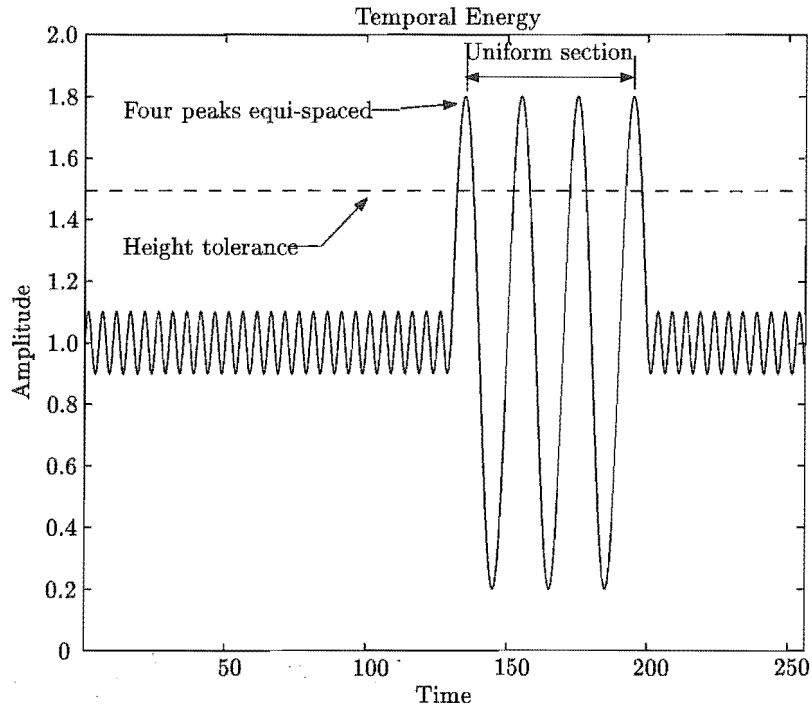


Figure 4.9 Schematic showing uniformity criteria based on the height and frequency of the temporal energy of a wavelet transform.

of each wavelet, the product of the four measures is taken. This produces the Wavelet Performance Measure, WPM, given by:

$$WPM = R_{harmonic} * R_{slope} * R_{CE} * R_{uniformity}. \tag{4.12}$$

The WPM is an open-ended scale that gives equal weighting to each component and to both frequency and temporal measures. For the ideal case and for the test signals used here, the Wavelet Performance Measure would be 10,000, indicating the least interference by scale harmonics, the best coincidence across scales and the best localisation in both the time and frequency domains.

4.5 PERFORMANCE OF WAVELETS

The Wavelet Performance Measure, WPM, is used to rank wavelets of the following families: biorthogonal daubechies (biorN_1.N_2), coifman (coifN), and symlet (symN), where N is the order of the wavelet. These wavelets are linear or nearly linear and are therefore expected to maintain coincidence across scales.

The signals used for assessment are 500 seconds in duration, sampled at 10 Hz. The frequencies chosen for the test signals are representative of frequencies found in infant breathing signals (0.03 - 1.3 Hz [Rantonen *et al.* 1998]).

The signal used for the first two assessment criteria of the wavelet performance measure, the rate of decay of scale harmonics and the slope of scalar energy, $R_{harmonic}$ and R_{slope}, Equations (4.5) and (4.7), consists of a single frequency at 0.5 Hz.

The signal used for the third assessment criterion, cumulative energy, R_{CE}, Equation (4.10), is zero with an impulse situated in the center of the signal. The cumulative energy is calculated for 10, 20, and 50 second intervals on either side of the impulse. The interval of 10 seconds is used for the rankings and is chosen because the concentration of energy is close to the impulse whilst still containing at least 50% of the cumulative energy for most signals.

The signal used for the fourth assessment criterion, uniformity, $R_{uniformity}$, Equation (4.11), consists of two frequencies. The first frequency of 0.5 Hz occurs during the first half of the signal and the second frequency of 0.05 Hz occurs during the second half of the signal. The transition is continuous in the zeroth order. The uniformity criterion is calculated for those wavelets that isolate only two frequency components. The period between peaks must be within 1% of the average period for those peaks to be included in the uniform section. The wavelets bior2.4, bior3.5, coif1, sym2 and bior1.1 have more than two frequency components and their uniformity ranking is set to 90, which is less than the uniformity ranking of the other wavelets tested.

The WPM is shown for scale-limited wavelets in Table 4.1, sorted by the highest overall ranking (best) wavelet at the top. Also shown in Table 4.1 are the number of vanishing moments, N, and the uncertainty principle bound for each wavelet, $\Delta t \Delta f$, where $(\Delta t)^2$ is the variance of the power distribution of the wavelet and $(\Delta f)^2$ is the variance of the power spectrum of the Fourier transform of the wavelet [Bracewell 1986]. An estimate of the uncertainty principle lower bound is calculated using an approximation of the wavelet generated to the 10th iteration of the filter coefficients.

Each wavelet was applied to the same typical quiet sleep infant breathing signal and the results in the wavelet domain were visually compared to the rank obtained by the wavelet performance measure. The performance of each wavelet showed good correspondence with the wavelet performance measure.

The performance of different wavelets is illustrated in two examples, with two wavelets applied to the typical quiet sleep infant breathing signal shown in Figure 4.10(a). The breathing signal exhibits two of the characteristics discussed previously: a main breathing frequency, and amplitude modulation. The wavelets used are the best wavelet according to the WPM, sym8, and the wavelet with the best uncertainty principle lower bound, bior4.4, shown in Figure 4.10(b) and 4.10(c), respectively. The scalar energy of each continuous wavelet transform is shown to its right. Both peaks in the sym8 transform (D1 and D2), Figure 4.10(d), have a higher scalar energy than the corresponding peaks in the bior4.4 transform (E1 and E2), Figure 4.10(e). The bior4.4 transform exhibits harmonics that can be seen in the continuous wavelet transform.

	New Performance Measures					Traditional Performance Measures	
Wavelets	Overall performance	Rate of decay	Maximum Slope	Cumulative Energy	Uniformity,	Vanishing moments	Uncertainty Principle lower bound
ψ	WPN	R_{harm}	R_{slope}	R_{CE}	R_{unif}	N	$\Delta t \Delta f$
sym8	2274	0.998	0.398	58.3	98.2	16	0.380
coif3	2233	0.995	0.376	61.8	96.6	6	0.356
bior6.8	1982	0.994	0.35	59	96.6	5	0.366
sym6	1944	0.993	0.368	54.2	98.2	12	0.356
bior2.8	1706	0.993	0.298	59.7	96.6	1	0.431
bior5.5	1202	0.496	0.399	62.9	96.6	4	0.338
sym7	1120	0.996	0.384	29.9	98	14	0.378
coif2	1040	0.497	0.335	64.7	96.6	4	0.338
bior2.6	835	0.496	0.283	61.6	96.6	1	0.395
bior3.9	737	0.497	0.346	43.7	98.2	2	0.433
sym4	735	0.497	0.328	46.7	96.6	8	0.341
bior4.4	700	0.333	0.338	64.4	96.6	3	0.320
sym3	589	0.331	0.303	61.9	95	6	0.412
bior1.5	557	0.499	0.238	47.8	98.2	0	5.759
bior2.4	504	0.332	0.265	63.7	90	1	0.350
bior3.7	493	0.333	0.334	45.9	96.6	2	0.400
sym5	485	0.499	0.35	28.8	96.6	10	0.373
bior3.5	472	0.332	0.322	49.1	90	2	0.359
coif1	413	0.248	0.268	69.2	90	2	0.530
bior1.3	358	0.332	0.223	51	95	0	5.113
sym2	110	0.143	0.264	32.5	90	4	0.594
bior1.1	59	0.054	0.204	59.6	90	0	4.903

Table 4.1 Rankings for wavelets based on traditional and novel performance measures. The wavelets are ordered by the overall performance measured by novel criteria. The best wavelet for all measures, except the uncertainty principle, is given by the greatest value for each measure. The best wavelet as measured by the uncertainty principle has the lowest value.

The first harmonic can be seen between scales 50 and 68 and a second harmonic can be seen between scales 90 and 110. The first harmonic is seen in the scalar energy at E3 (Figure 4.10(e)) as a minor peak at a scale of 59. The second harmonic is seen in the scalar energy at E4 (Figure 4.10(e)) as a small peak on the side of peak E2 at a scale of 98. Despite the bior4.4 wavelet having a better uncertainty principle lower bound than the sym8 wavelet, the bands due to individual frequency components in the sym8 transform (Figure 4.10(b)) are more distinct than the bands in the bior4.4 transform (Figure 4.10(c)). The lower extent of the low frequency peak (E2) is difficult to establish in the bior4.4 continuous wavelet transform or scalar energy representation, due to the presence of the harmonic, but it appears to be similar to the sym8. However, the upper extent of the low frequency peak (D2) is at scale 225 for the sym8 wavelet (Figure 4.10(d)), while the upper limit of the low frequency peak (E2) is at scale 250 for the bior4.4 wavelet (Figure 4.10(e)). Using the sym8 wavelet, the discrimination between the main breathing frequency and the amplitude modulation is improved, confirming the accuracy of the WPM as a measure of performance.

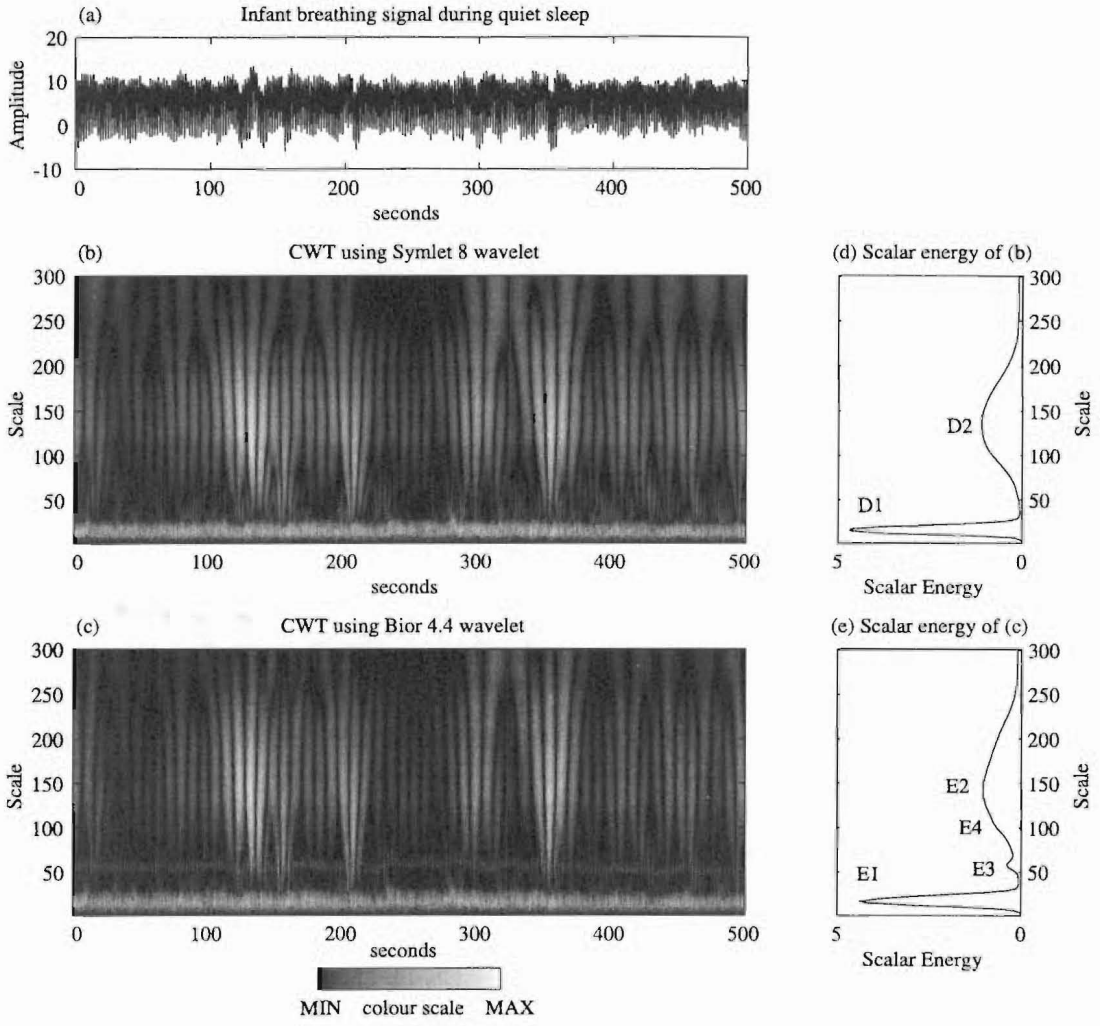


Figure 4.10 Examples of analysing an infant breathing signal from quiet sleep using two different wavelets. (a) Infant breathing signal. (b) Continuous Wavelet Transform of (a) using Symlet 8 wavelet. (c) Continuous Wavelet Transform of (a) using Bior4.4 wavelet. (d) Scalar energy of (b), with local maxima at D1 and D2. (e) Scalar energy of (c), with local maxima at E1 to E4.

4.6 DISCUSSION

The results of applying the Wavelet Performance Measure, WPM, presented in Table 4.1 indicate that the most appropriate wavelet to analyse infant breathing signals is the symlet of order 8. This is somewhat surprising as, originally, it was thought that a linear phase wavelet would be the best choice for analysis since these wavelets are generally regarded as having the least distortion in the wavelet domain [Vetterli and Kovačević 1995]. In the two temporal tests, in which the linear wavelets were expected to perform well, either symlet or coiflet wavelets performed better or equivalently to the biorthogonal wavelets. Even between the biorthogonal wavelet families, there is a large variation in performance based on the WPM.

Higher order symlets, in general, perform better than lower order symlets. How-

ever, even order symlets perform better than the odd order symlet one order higher. The reason for this is unclear, but may be related to the symmetry of the scaling functions and the wavelet functions as the even order symlets tend to be more symmetrical than the odd order symlets.

The symlet wavelet of order 8 does have the highest number of vanishing moments of the wavelets tested, but, as can be seen in Table 4.1, the number of vanishing moments is by no means the best indicator for choice of wavelet when compared with the Wavelet Performance Measure. Originally it was thought that a higher number of vanishing moments would indicate better performance. Indeed, within any given wavelet family, the higher the number of vanishing moments, the better the wavelet performed. However, this linear relationship is lost comparing wavelets between families, according to the Wavelet Performance Measure.

The performance of the wavelets given by the uncertainty principle lower bound does not relate directly to either the number of vanishing moments or the wavelet performance measure. Even though the uncertainty principle is used to select wavelets [Unser and Aldroubi 1996], the results of the wavelet performance measure show that, for analysing amplitude modulation of infant breathing, it may not be an appropriate selection criterion.

The Wavelet Performance Measure for wavelets analysing a particular type of signal produces different results than traditional approaches. This is because traditional methods use only the characteristics of the wavelets in isolation to determine the ranking and do not incorporate the salient characteristics of the signal to be analysed. The results in Figure 4.10 show that the sym8 wavelet, selected with the new method, does give a clearer representation of the breathing signal in the wavelet domain compared with the bior4.4 wavelet, selected with a traditional method. The higher scalar energy of the peaks, absence of harmonics, and narrower bands for each frequency component make it easier to isolate different frequency components in the sym8 continuous wavelet domain. This finding supports the use of a technique for selecting a wavelet that is specific to the problem at hand, such as using the Wavelet Performance Measure to select a wavelet specifically for analysing breathing.

The Wavelet Performance Measure is the direct product of four components, and each component is given equal weighting (Equation (4.12)). If an application warranted different weighting for temporal and frequency components, for example, then a ratio of the two temporal component ranks and the two frequency component ranks could be used with one group weighted more heavily.

The Wavelet Performance Measure used to select a wavelet for analysing infant breathing signals is designed to isolate the pertinent characteristics found in quiet sleep. However, a wavelet selected in this manner may not be appropriate for analysing breathing recorded during other sleep states that have substantially different characteristics,

such as the highly variable breathing rate of rapid eye movement sleep.

4.7 CONCLUSIONS

Two main features of continuous wavelet transforms applied to signals are shown in the wavelet domain by analysing test signals. Firstly, signals with a frequency component that exists for a short time exhibit a fan-like pattern in the continuous wavelet domain. Secondly, signals with a frequency component that exists for a long time exhibit vertical stripes in the continuous wavelet domain at a frequency proportional to the frequency in the signal. Signals that contain a combination of frequency components existing for short and long times exhibit a combination of vertical stripe and fan-like patterns in the continuous wavelet domain.

The Wavelet Performance Measure, WPM, is presented as a novel method for selecting a wavelet for the analysis of infant breathing signals recorded during quiet sleep. The WPM is developed to address the problem that theoretical methods for selecting wavelets may not provide sufficient information for specific analysis situations. The WPM uses four measures based on the energy of the wavelet transform of the test signals to determine the overall performance of the wavelet. The four measures test the ability of the wavelets to localise both frequency and temporal information in the wavelet domain. Test signals that model various features of interest in the respiration signals are used. Using the Wavelet Performance Measure, of the wavelets tested, the best is the Daubechies symlet wavelet of order 8. When applied to breathing signals, the symlet wavelet of order 8 exhibits more desirable properties in the wavelet domain than wavelets selected using traditional methods. The new method selects a wavelet that improves isolation of features of interest, and suggests that in other applications, a refined selection method for the analysing wavelet may result in superior performance.

Chapter 5

INFANT BREATHING SIGNAL ANALYSIS USING WAVELETS

This chapter presents the analysis of infant breathing using the wavelet transform. Characteristics of breathing signals in the wavelet domain are described and two measures are developed to quantify aspects of one of those characteristics, namely amplitude modulation. The results of analysis using the two measures are presented. The analysis had two main aims: 1) to measure amplitude modulation of breathing during quiet sleep and 2) to evaluate differences in breathing patterns between infants at high and low risk for Sudden Infant Death Syndrome (SIDS).

The first section describes a preliminary study of infant breathing, showing patterns in the wavelet domain for different breathing characteristics and indicating the direction for further research. The Fourier Transform and the Short Time Fourier Transform are compared to the wavelet transform for isolating frequency components in the second section. A new measure, *rhythmicity*, is defined to assess the extent of constant frequency components in infant breathing during quiet sleep. Rhythmicity presents an alternative to spectral analysis based on the Fourier Transform used in other studies [Rantonen *et al.* 1998] for analysing low frequency components of infant breathing signals. A second new measure, *impact*, is defined to assess the influence that any one frequency component has on the signal. These two measures, rhythmicity and impact are presented in the third section. The fourth section describes assessment of rhythmicity for infants who were at high risk of Sudden Infant Death Syndrome (SIDS) compared with infants who were at low risk of the syndrome. The fifth section describes an initial study that compares the impact of amplitude modulation in breathing from infants who later succumbed to SIDS with normal infants. This is followed by a more extensive study investigating time of night effects upon amplitude modulation of breathing from infants who later succumbed to SIDS and their controls. Conclusions are presented in the final section.

5.1 CHARACTERISTICS OF BREATHING SIGNALS IN THE WAVELET DOMAIN

This section describes a preliminary analysis performed on data from infants at high risk for Sudden Infant Death Syndrome (SIDS) and infants at low risk for the syndrome. One aim is to visually assess wavelet transforms of breathing from various sleep states and to determine if the frequency components of the principal breathing frequency and amplitude modulation are isolated in the wavelet transform domain. It was expected that the amplitude modulation would appear at higher scales than the primary breathing frequency. A second aim was to determine whether there was a difference in the characteristics of wavelet transforms from breathing signals between infants at high risk for SIDS and infants at low risk for SIDS. The outcome of this analysis directed the development of measures for assessing pertinent characteristics found in the wavelet domain, described in later sections.

5.1.1 Methods

A preliminary analysis was performed using a set of data recorded from sleeping infants in their home environment. Twenty nights of breathing were recorded from 14 infants [Brown *et al.* 1992, Tappin *et al.* 1996a]. The infants ranged in age from 48 days to 80 days. Ten of the recordings were from four infants at high risk for SIDS. These infants were classified as high risk because they had suffered an Apparent Life Threatening Event (ALTE) [Brown *et al.* 1992]. The remaining ten recordings were from ten infants at low risk for SIDS. These infants were classified as having a low risk for SIDS because the following criteria were met: non-smoking household, predominantly breastfed, slept on their backs or their sides, parents living together, parents older than 20 years old, non-Maori,[1] not the first born in the family, no siblings that have died of SIDS, full term delivery, no serious illness or history of breathing problems, and a birth weight between 3000 and 4300 grams [Tappin *et al.* 1996a].

All 20 recordings were analysed for frequency and locational features using wavelets, and, in particular, for characteristics of amplitude modulation in the wavelet domain that could lead to distinguishing between high risk and low risk infants. The Daubechies symlet wavelet of order 8 was used in the analysis, selected according to the Wavelet Performance Measure presented in Chapter 4 [Macey and Page 2000]. At least six sections of breathing were analysed from each night of breathing. The sleep state was classified as undetermined, Rapid Eye Movement (REM) sleep, or quiet sleep (refer to Section 1.2). The extent of amplitude modulation in the wavelet transform was assessed for each section.

[1] Maori are the indigenous people of New Zealand.

5.1.2 Results

Amplitude modulation is exhibited by both risk groups, but is more evident in the group of infants at high risk for SIDS during quiet sleep. This trend is shown in Table 5.1 where the wavelet transforms of 60 epochs of sleep from each group were classified as having greater than 50% of the duration of the epoch with amplitude modulation, less than 50% amplitude modulation but more than 10% amplitude modulation, or as having less than 10% amplitude modulation. Sleep state was either undetermined, quiet or REM. A total of 12 epochs showed greater than 50% amplitude modulation for the infants at high risk for SIDS, and only 5 epochs showed greater than 50% amplitude modulation for infants at low risk for SIDS. Conversely, only 27 of the epochs showed less than 10% amplitude modulation for infants at high risk for SIDS, compared with 38 epochs for infants at low risk for SIDS.

Risk for SIDS	Type of sleep	Amplitude Modulation $> 50\%$	Amplitude Modulation $< 50\%, > 10\%$	Amplitude Modulation $< 10\%$
High	Undetermined	7	14	19
	Quiet sleep	5	6	1
	REM sleep	0	1	7
Low	Undetermined	3	13	26
	Quiet sleep	2	3	4
	REM sleep	0	1	8

Table 5.1 Number of epochs with varying percentages of the epoch duration with amplitude modulation for different sleep states for infants at high and low risk for SIDS.

These findings are illustrated with a number of examples, each consisting of a section of breathing lasting eight minutes 20 seconds (500 seconds or 5000 samples). The continuous wavelet transform was calculated for either 512 scales or 1536 scales. The examples have had their edge effects removed. Edge effects act like impulse responses giving a fan-like structure, as shown in Figure 4.1. For the analyses calculated to 512 scales, the wavelet transform was calculated for 8000 samples and only the middle 5000 shown. Truncating the wavelet transform removes the edge effects. For the analyses calculated to 1536 scales, the wavelet transform was calculated for 10000 samples with the middle 5000 shown. Truncating the transforms removes the edge effects up to scales of approximately 1000, depending on the intensity in the main part of the wavelet transform. The remaining edge effects are insignificant compared with the detail shown in the main part of the continuous wavelet transform.

The first example, shown in Figure 5.1, shows a prolonged section of periodic breathing from an ALTE infant. None of the normal infants exhibited such large amplitude modulation for this period of time. The continuous wavelet transform shows two main frequencies present in the section of breathing. The higher frequency, at a scale of 20, is the frequency of the infant's breath-to-breath intervals, interspersed with

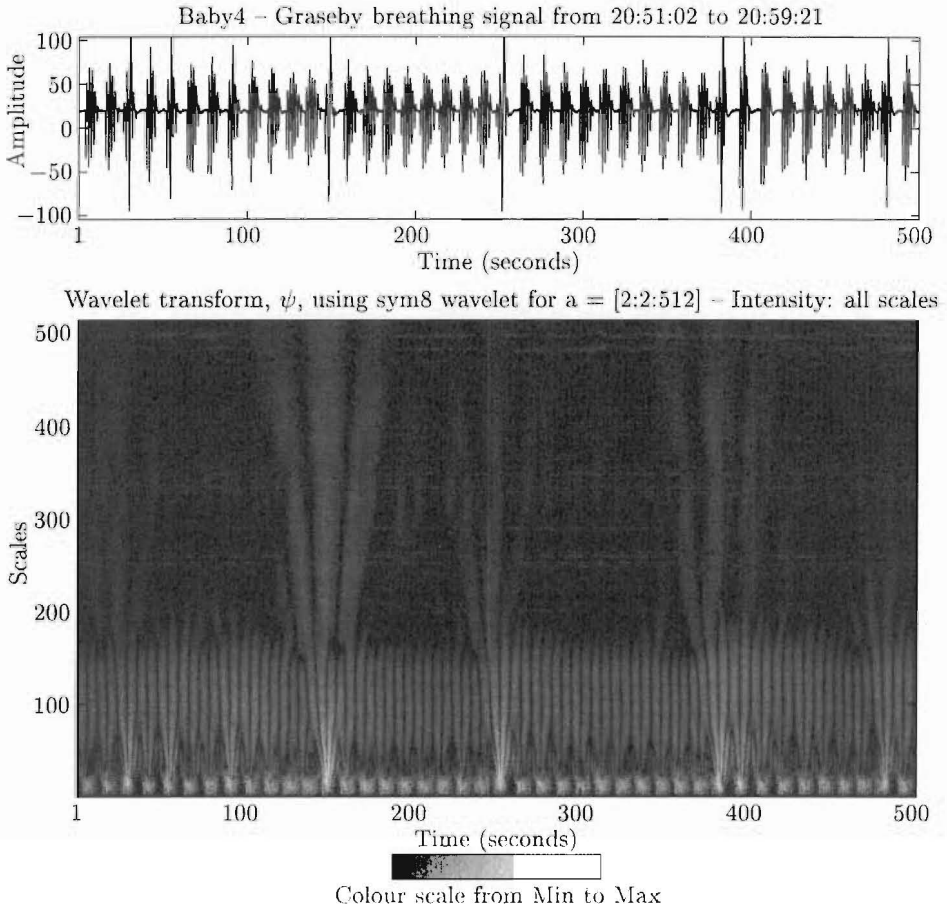


Figure 5.1 ALTE infant's breathing (top) and continuous wavelet transform (bottom).

apnoea. There is also a band of the transform that is related to the frequency of the apnoea: at a scale of 100, the frequency is 0.096 Hz, which is equivalent to a period of 10.4 seconds. None of the low risk infants in this group exhibited breathing with regular apnoea.

The second example, given in Figure 5.2, shows a section of breathing from a low risk infant. The continuous wavelet transform shows almost no periodicity in the signal at any scale. The lack of any consistent patterns indicates that there are no sustained frequency components in this example. This is in contrast to the two frequency components isolated in the previous example in Figure 5.1. None of the ALTE infants exhibited breathing with so little energy in the continuous wavelet domain for scales above 50.

The third and fourth examples are of very regular breathing from periods of quiet sleep. The third example exhibits mild amplitude modulation, while the fourth example shows no obvious periodicities in the breathing signal apart from the actual breathing rate.

The third example in Figure 5.3 is from a second ALTE infant. It was rare to

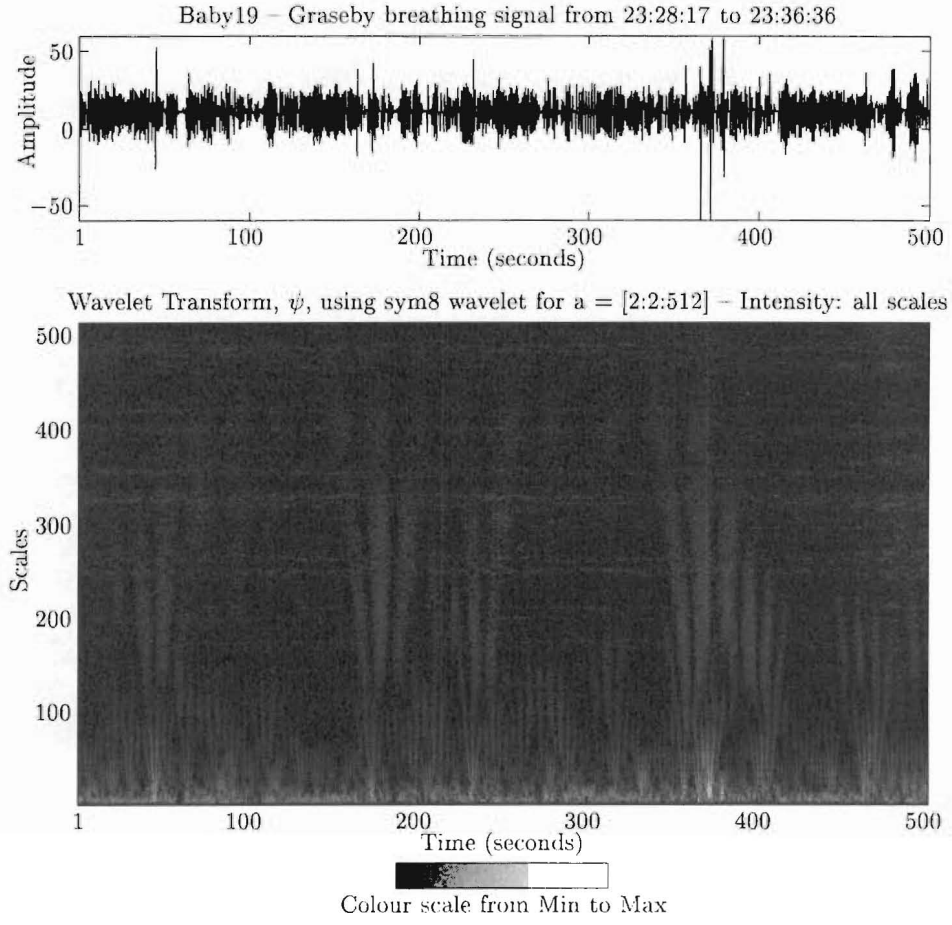


Figure 5.2 Low risk infant's breathing (top) and continuous wavelet transform (bottom).

find such a long period of sleep with no apnoea in the ALTE infants. As with the first example, the principal breathing frequency is present in the continuous wavelet transform at a scale of 20. A second periodicity is also present at a scale of 150, representing a frequency of 0.057 Hz, and hence a period of 17.5 seconds. A third frequency is present during the last 200 seconds of the section, represented by regular vertical bands at a scale of 400. This scale corresponds to a frequency of 0.02 Hz and period of 50 seconds. This example of breathing shows, therefore, two frequencies modulating the breathing. While the modulation at 0.057 Hz can be visually discerned in the breathing signal, the second modulation frequency at 0.02 Hz cannot be easily seen.

The fourth example in Figure 5.4 is from a second low risk infant. The breathing signal has a similar magnitude to the third example and has no obvious periodicities apart from the principal breathing rhythm. The continuous wavelet transform shows the frequency of the breathing at a scale of about 20, which is similar to the previous example. Unlike the first low risk infant in Figure 5.2, this infant's continuous wavelet transform does show some patterns at higher scales. However, unlike the ALTE infants

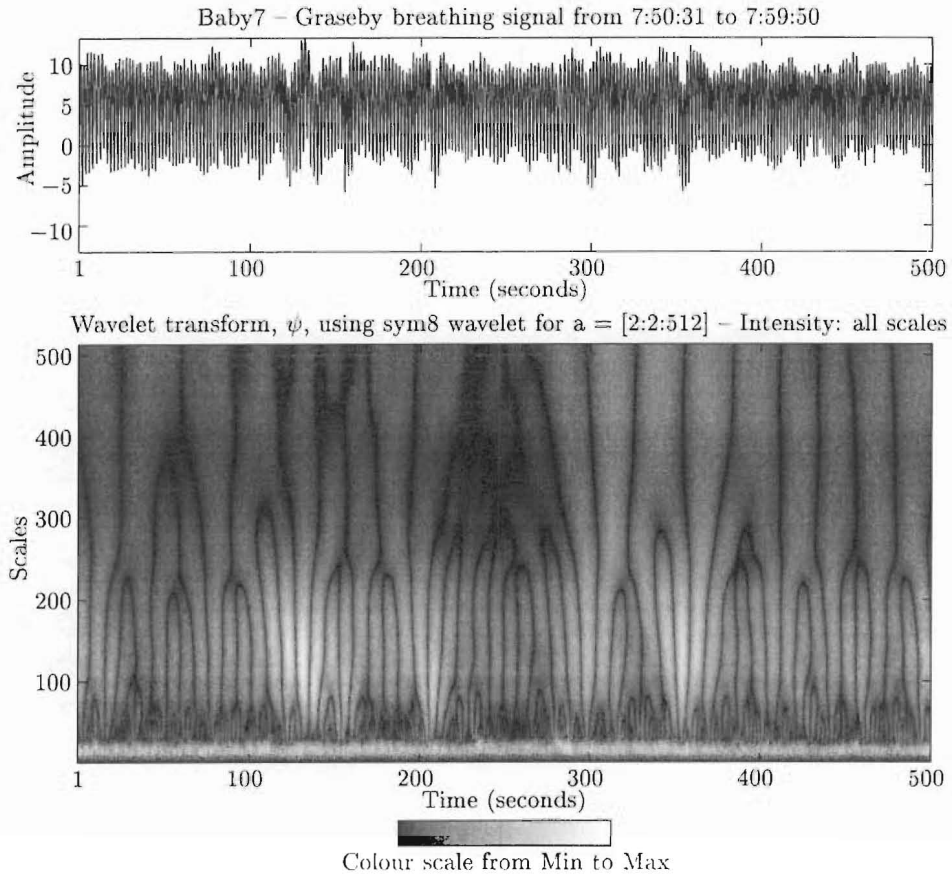


Figure 5.3 ALTE infant's breathing (top) and continuous wavelet transform (bottom).

in Figures 5.1 and 5.3, there is no consistency to the patterns. The patterns are more fan-like than like vertical strips, suggesting that the patterns result from delta functions or continually changing frequency components (see Section 4.2). The fan-like patterns correlate with the large sighs evident in the original breathing signal.

The fifth example in Figure 5.5 is from a third ALTE infant. The continuous wavelet transform is to 1536 scales and shows a periodicity at a scale of 600. The frequency of this periodicity is 0.016 Hz, with a period of 62.5 seconds representing amplitude modulation of the breathing.

These examples illustrate a range of breathing characteristics and their corresponding characteristics in the wavelet domain, as summarised in Table 5.1. Constant frequency components appear as vertical stripes in the wavelet domain; very short time transient components appear in the wavelet domain as fan-like patterns. For the signals analysed using the sym8 wavelet, the principal breathing frequency was seen at a scale about 20. Amplitude modulation is seen at higher scales (100 - 800), isolated from the principal breathing frequency.

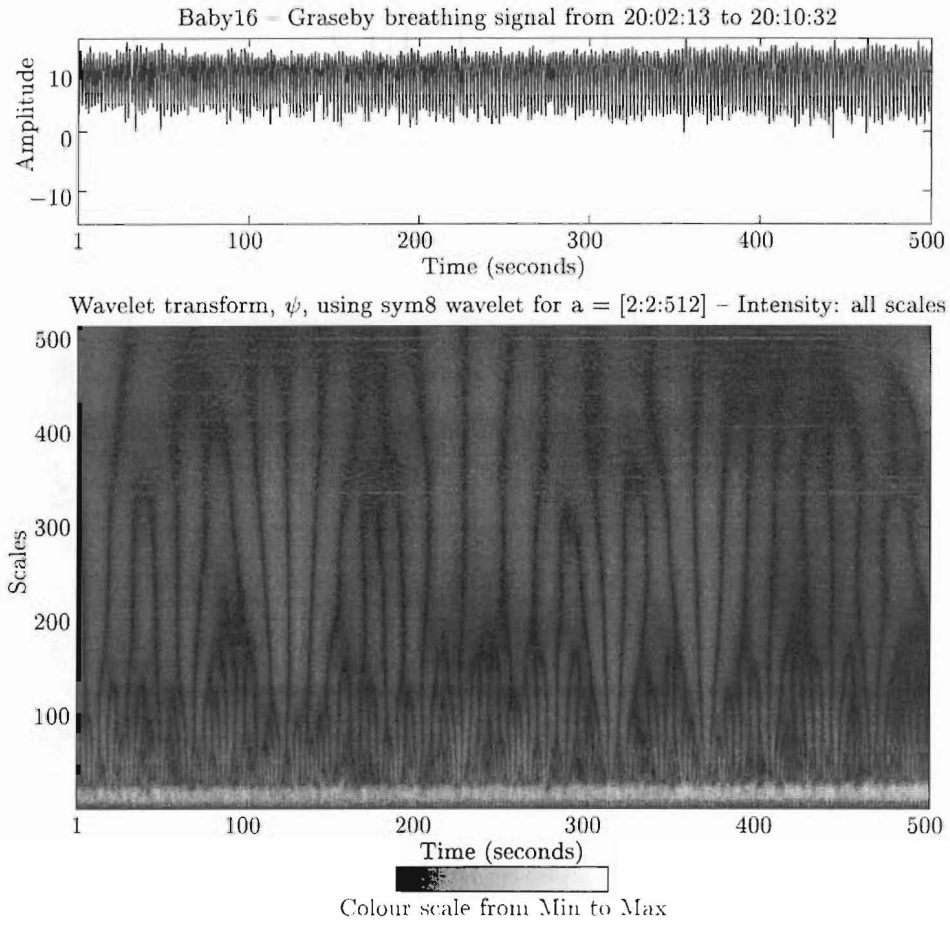


Figure 5.4 Low risk infant's breathing (top) and continuous wavelet transform (bottom).

5.1.3 Discussion

Breathing within a sleep state exhibited consistent characteristics in the wavelet domain, but the characteristics differed between states. Breathing during quiet sleep showed more constant frequencies, including amplitude modulation, than breathing during other states of sleep. The difference between states is expected, as it has been shown with other measures [Schechtman *et al.* 1988]. However, amplitude modulation has not previously been explicitly measured. In this initial study, it was found that amplitude modulation occurs preferentially in quiet sleep. Furthermore, the ALTE infants, at high risk for SIDS, appeared to show more amplitude modulation during quiet sleep than infants at low risk for SIDS, consistent with the finding by Rantonen *et al.* [1998].

Breathing that is erratic can be identified by the presence of fan-like patterns in the wavelet domain. The fan-like patterns are similar to the patterns obtained using the delta function in the test signals. The breathing rate of regular breathing can be identified by vertical stripes at a low scale in the wavelet domain. The vertical stripes are similar to the pattern obtained using sinusoids for the test pattern. The wavelet

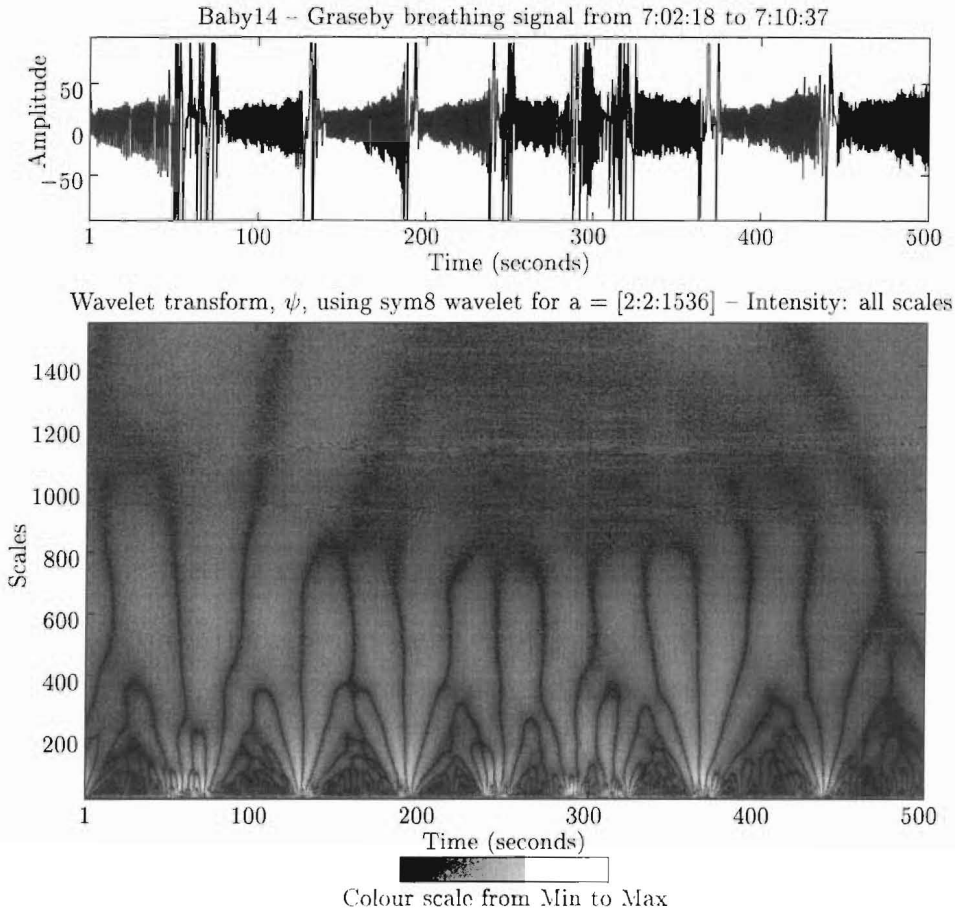


Figure 5.5 ALTE infant's breathing (top) and continuous wavelet transform (bottom).

transforms of the delta function and sinusoid test patterns assisted in the interpretation of these results.

Amplitude modulation is identified by vertical stripes in the wavelet domain at higher scales that are isolated from the scales representing the principal breathing rate. The isolation of the amplitude modulation in the wavelet domain is necessary for the assessment of the impact of those frequency components on the remainder of the signal. Impact of a frequency component on the signal is determined by the magnitude, duration and consistency of the frequency of that component. For a given frequency, the larger the magnitude, longer the duration, or more consistent the frequency, the larger the impact of a particular frequency component on the signal.

It is possible that simpler methods of frequency analysis would also measure amplitude modulation. Spectral analysis has been used to quantify the ratio of low to high frequency signal components [Rantonen *et al.* 1998] in infant breathing. However, as seen in this section, frequency components in breathing are non-stationary, therefore, it is likely a non-stationary analysis is more appropriate. This hypothesis is tested in the next section.

5.2 COMPARISON OF BREATHING SIGNAL CHARACTERISTICS IN FOURIER AND SHORT TIME FOURIER DOMAINS

In this section, characteristics of breathing signals in the wavelet domain are compared to characteristics obtained using the Fourier Transform and the Short Time Fourier Transform. These two types of Fourier Transforms were applied to breathing data analysed by the wavelet transform and the results were compared, with respect to extraction of amplitude modulation frequency and location in the signal. Based on the theoretical background that Fourier transforms assume a stationary signal, and that the wavelet transform exploits the Heisenburg Uncertainty Principle better than the Short Time Fourier Transform, the wavelet transform would be expected to perform more appropriately for a non-stationary signal. In testing, the wavelet transform was shown to distinguish the frequency and locational properties of amplitude modulation better than the Fourier Transform based methods.

Figure 5.6 shows the results of one example of breathing analysed using both the Fourier Transform and the Short Time Fourier Transform. This result is indicative of the comparisons made between the Fourier techniques and the wavelet transform. The breathing signal used for this comparison is the signal shown at the top of Figure 5.6. It shows an eight minute, 20 second period of quiet sleep from an infant at high risk for SIDS. The Fourier Transform is shown in the center of Figure 5.6. A high peak can be seen at 0.5 Hz, which represents the infant's principal breathing rate. Two smaller peaks can be seen at 0.057 Hz and at 1 Hz. The peak at 1 Hz corresponds to the frequency of the infant's heart beat, and the peak at 0.057 Hz corresponds to one frequency component of the amplitude modulation isolated using wavelet transforms. However, in the analysis using wavelet transforms, another frequency component at 0.02 Hz was isolated. This frequency component can be seen as a small peak in the Fourier Transform, but is not represented by a major feature and would be difficult to objectively distinguish from the remaining frequency components. The significance of this peak is less in the Fourier Transform because the frequency component was of small amplitude and only present in the last two fifths of the signal. This example highlights the limitations of assuming a stationary signal, as is the case when using the Fourier Transform.

The Short Time Fourier Transform, shown at the bottom of Figure 5.6, was calculated using a window length of 512, providing good resolution in the frequency domain and moderate resolution in the time domain. The main breathing rate can be seen at 0.5 Hz as a bright band across the plot. The infant's heart rate is visible at 1 Hz. However, the remaining frequency components at 0.057 and 0.02 Hz that were distinguished using the wavelet transform cannot be observed.

Although the Fourier Transform and the Short Time Fourier Transform may con-

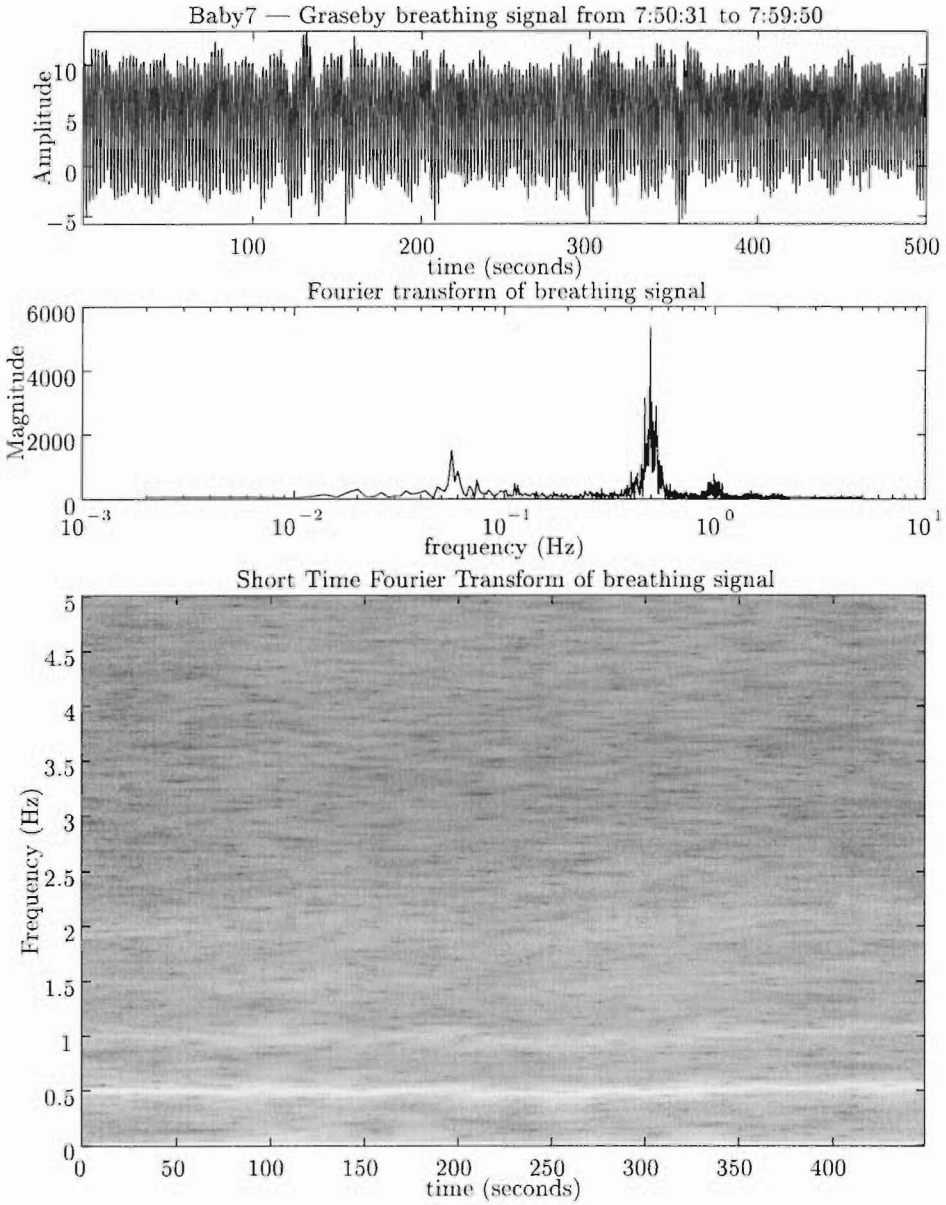


Figure 5.6 Breathing signal (top), Fourier Transform of Signal (middle), and Short Time Fourier Transform (bottom).

tain information regarding frequencies of interest, all frequencies and their durations are more clearly distinguished using the continuous wavelet transform. This is especially relevant when designing an algorithm to objectively measure these features.

Several time-frequency analysis approaches have been used for biomedical signals, including Short Time Fourier Transforms or variations [Ruhsam *et al.* 1995, Keselbrener and Akselrod 1996, Dehkordi *et al.* 1999], Wigner-Ville distributions [Ruhsam *et al.* 1995, Chan *et al.* 1997, Dehkordi *et al.* 1999] and matching pursuit, a variation of wavelet transforms [Akay and Mulder 1996]. These approaches have been successfully applied to cardiac signals and for isolating breathing and cardiac components of signals. It is possible that these methods would also allow measurement of amplitude modulation and that they would also perform better than the Fourier Transform.

Amplitude modulation is isolated more successfully in the wavelet transform compared with the Fourier Transform or Short Time Fourier Transform. Initial differences between infants at high and low risk for SIDS, presented in Section 5.1, also provided motivation for developing a novel measure, rhythmicity, based on the wavelet transform, to assess the extent of significant, constant frequency components in a signal. Rhythmicity is subsequently applied to the breathing signals of both ALTE infants and SIDS infants and compared with results for control infants in Sections 5.4 to 5.6.

5.3 RHYTHMICITY AND IMPACT: NOVEL MEASURES OF FREQUENCY COMPONENTS

A primary aim of studying breathing of infants at high risk for SIDS, or who later succumb to SIDS, and normal infants is to provide measurable characteristics of breathing that differentiate the two groups. As discussed in Section 4.1, the pertinent characteristics of infant breathing during quiet sleep are: a primary breathing frequency with low variability in amplitude or frequency relative to other sleep states, amplitude modulation of various magnitudes of the primary breathing frequency, and large inspiratory efforts followed by pauses in breathing. Amplitude modulation and the effect it has on the primary breathing frequency are the focus of this chapter, as there has been little previous research into this phenomenon.

To assess the extent of amplitude modulation, the concept of *rhythmicity* is introduced. Rhythmicity is the extent to which frequency components remain constant for the duration of the breathing signal [Macey *et al.* 2000a]. Rhythmicity may be calculated for individual frequency components or for the entire signal. For example, the influence of amplitude modulation on the primary breathing frequency can be assessed [Macey *et al.* 2000b]—the magnitude of this influence is defined by *impact*. The *overall* rhythmicity of a signal is defined as the percentage of total energy in uniform sections of the continuous wavelet transform.

Sections in the wavelet domain are defined as *uniform* where the energy of the continuous wavelet transform has significant magnitude and constant frequency. Two aspects of the energy are evaluated: 1) the energy at each scale across all time, *scalar energy*, which identifies the distribution of the energy across scales, Equation (4.3), and 2) the energy at each time across scales, *temporal energy*, Equation (4.4). The uniform sections are identified in five steps:

1. The scales that relate to the significant frequencies in a signal are identified by peaks in the scalar energy. The peaks in the scalar energy are found using the derivative of the scalar energy. A peak, E_P, is identified at scale j where $\frac{dE_{\phi,t}(j)}{dj} = 0$ and $\frac{dE_{\phi,t}(j_-)}{dj} > 0$ and $\frac{dE_{\phi,t}(j_+)}{dj} < 0$, where j_- and j_+ are just less than j and more than j respectively. The highest peak, $E_{P_{max}}$ is identified as the peak due to the primary breathing frequency. The remaining peaks are considered significant frequency components if their height is at least 1% of the height of $E_{P_{max}}$, i.e., $E_{P_{sig}} = E_P$ where $E_P > 0.01E_{P_{max}}$.

2. A range of scales encompassing each peak in scalar energy is determined. Each range of scales is bound on both sides by either a trough in the scalar energy or the scalar energy decreasing to a tolerance, whichever is closer to the peak. A trough in scalar energy is defined using the derivative of the scalar energy, in much the same way as the peaks are defined. A trough, E_T, is identified at scale j_T where $\frac{dE_{\phi,t}(j)}{dj} = 0$ and $\frac{dE_{\phi,t}(j_-)}{dj} < 0$ and $\frac{dE_{\phi,t}(j_+)}{dj} > 0$. The tolerance is defined as $E_{tol} = 0.01E_{P_{max}}$, occurring at scale j_{tol}. The range of scales is given by $j_{P_-} : j_{P_+}$ surrounding a peak at scale j_P such that if $j_{tol} < j_T < j_P$ then $j_{P_-} = j_T$, but if $j_T < j_{tol} < j_P$ then $j_{P_-} = j_{tol}$. Similarly, for j_{P_+}, if $j_P < j_T < j_{tol}$ then $j_{P_+} = j_T$, but if $j_P < j_{tol} < j_T$ then $j_{P_+} = j_{tol}$.

3. Temporal energy is calculated only for the range of scales surrounding each peak. Equation (4.4) is applied for each peak with limits set to the corresponding range of scales, $j_{-P} : j_{+P}$.

4. The frequency of the wavelet transform, and therefore the signal, is calculated for the temporal energy bands using the major peaks in the temporal energy. The major peaks are found using a peak detection algorithm [Tuffnell 1993] and the period, T, between peaks is used to calculate frequency, f, in the normal manner ($f = \frac{1}{T}$).

5. Sections are defined as uniform where peaks in the temporal energy bands are close in frequency and a magnitude threshold is exceeded. Peaks are defined as close in frequency where consecutive peaks (at least three) have a period within 30% of the average period, $\bar{T}$, for the particular temporal energy band. If the time at which a peak occurs is t_P, and consecutive peaks are at $t_{P,i}$, and $t_{P,i+1}$, then the periods between the peaks are given by $T_i = t_{P,i+1} - t_{P,i}$. If T_i and

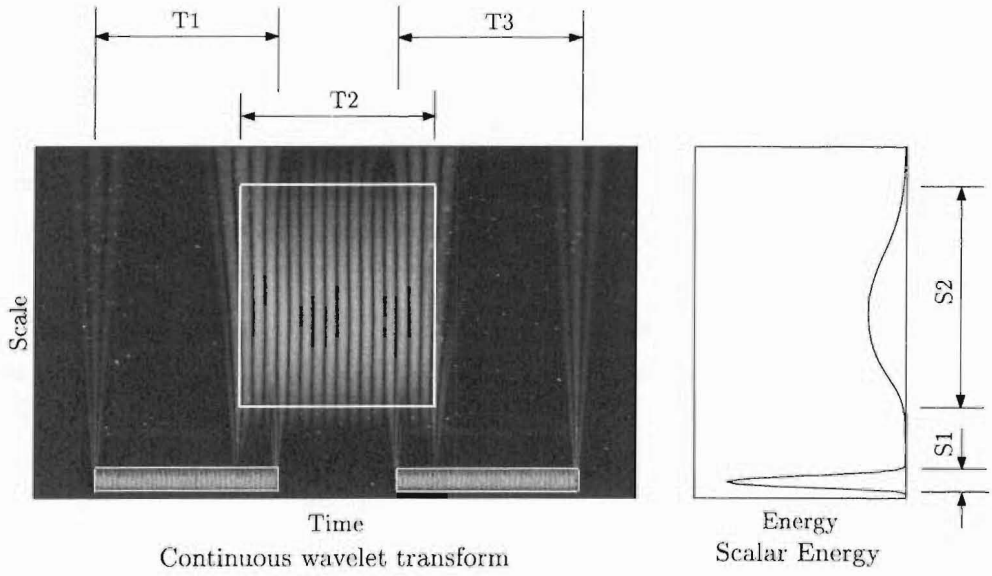


Figure 5.7 Continuous wavelet transform and scalar energy of test signal. The test signal is 5000 samples and has one frequency component of 0.5 Hz in two sections each lasting for a third of the signal and a second frequency component of 0.05 Hz in the middle third of the signal. In the continuous wavelet transform, black represents very low values. The vertical bands represent areas that peaks in scalar and temporal energy have significant magnitude. Uniform sections of continuous wavelet transform are contained within scales S1 and S2 and times T1, T2 and T3 (outlined in white.)

T_{i+1} satisfy $\frac{2\bar{T}}{3} \leq T_i \leq \frac{4\bar{T}}{3}$, then the section between $t_{P,i}$ and $t_{P,i+2}$ is defined as uniform. Successive peaks are tested to determine the boundaries of the uniform sections.

Once the uniform sections are identified, the sum of the energy in all the sections, E_{US}, is expressed as a percentage of the total energy, E_ψ, in the wavelet transform. This percentage is the rhythmicity, $R_{\psi,J}$, of the signal for a given wavelet, ψ, and number of scales, J:

$$R_{\psi,J} = \frac{100 E_{US}}{E_\psi}. \tag{5.1}$$

For example, in Figure 5.7 the continuous wavelet transform and scalar energy of a test signal is shown. To calculate the *overall* rhythmicity, the energy is summed for the scales S1 and S2 at times T1, T2 and T3.

For an *individual* frequency, rhythmicity describes the influence of that frequency whose energy is constrained to be locally uniform across scale and time. Thus, this measure can be used to assess the effect of modulating frequencies on the primary breathing frequency. A second measure, *impact*, is defined to quantify this effect. The impact, I, of amplitude modulation on the primary breathing frequency is expressed by the rhythmicity due to amplitude modulation, R_{AM}, as a percentage of the total

rhythmicity, R_T:

$$I = \frac{100 R_{AM}}{R_T}. \tag{5.2}$$

For example, to calculate the rhythmicity of the individual frequencies in Figure 5.7, the energy is summed for scale S1 at times T1 and T3 and, separately, for scale S2 at time T2.

Some additional aspects need to be considered when calculating rhythmicity. As noted earlier, the energy of a continuous wavelet transform for a given signal depends on the particular wavelet used in the analysis, as well as the number of scales calculated [Odegard *et al.* 1992]. In order to compare the effect of amplitude modulation between breathing signals, these parameters must be kept constant for all analyses. Boundary effects must also be taken into consideration by ensuring that the energy associated with the boundary effects is significantly less than the energy of the signal for the wavelet transform. For the signals analysed in Sections 5.4 to 5.6, the method used to reduce the boundary effects was symmetric extension, which has been shown to effectively reduce boundary effects [Chen Li *et al.* 1995].

5.4 RHYTHMICITY OF BREATHING DURING QUIET SLEEP FROM ALTE AND CONTROL INFANTS

This section investigates the extent to which the primary breathing frequency is regular and the extent to which amplitude modulation is present in the breathing of quiet sleep of infants at high risk for Sudden Infant Death Syndrome (SIDS) and infants at low risk for SIDS. The new measure defined in Section 5.3, rhythmicity, is used to quantify the uniform sections of frequency components, and the results are compared for the two groups. This study is designed as a limited evaluation, with the aim of assessing the possible use of rhythmicity as a measure that could distinguish between the two groups. It is not intended as a comprehensive physiological study.

5.4.1 Methods

Breathing signals recorded from eight of the infants used in the preliminary study of breathing characteristics in the wavelet domain are analysed (see Section 5.1). The infants are evenly divided into two groups. The four infants that had suffered an ALTE were in one group and the second group were at low risk for SIDS.

Epochs of 500 seconds of sleep that were identified as occurring during uninterrupted quiet sleep were analysed for the eight infants. Epochs were selected that exhibited obvious amplitude modulation where possible. The Daubechies symlet wavelet of order 8 was used in the analysis, selected according to the Wavelet Performance

Measure presented in Chapter 4 [Macey and Page 2000]. The wavelet transform was calculated, starting at a scale of two, in steps of two, to a scale of 500.

Rhythmicity (Equation (5.1)) was calculated for each epoch to assess the effect of amplitude modulation on the primary breathing frequency. The overall rhythmicity, rhythmicity of the primary breathing frequency, and the rhythmicity of the amplitude modulation were calculated and compared for the low and high risk for SIDS groups.

5.4.2 Results

Rhythmicity was calculated for eight infants and the results shown in Table 5.2. Table 5.2 also shows the rhythmicity for the main breathing frequency and for any modulating frequencies. Both groups of infants show a large variation in the overall rhythmicity of their breathing signals. The infants at low risk for SIDS show little or no rhythmicity of the amplitude modulation. The infants at high risk for SIDS, however, all show some rhythmicity of the amplitude modulation. Two of the infants at high risk for SIDS show a large amount of rhythmicity of the amplitude modulation.

	SIDS risk	Overall Rhythmicity	Rhythmicity of main breathing frequency	Rhythmicity of modulating frequencies
	High	56.4	37.1	19.3
	High	78.1	75.6	2.5
	High	81.3	80.5	0.8
*	High	83.8	26.6	57.2
	Low	30.4	27.7	2.7
	Low	56.0	52.3	3.7
	Low	66.2	66.2	0.0
†	Low	97.4	97.4	0.0

Table 5.2 Rhythmicity calculated for eight infants.(* Result for example in Figure 5.8, † result for example in Figure 5.9.)

The breathing signal for the infant at high risk for SIDS showing the greatest amount of overall rhythmicity is shown in Figure 5.8(a) with corresponding continuous wavelet transform in Figure 5.8(b). The wavelet transform is shown to 300 scales only. There were no significant features above 300 scales. The main breathing frequency of 0.5 Hz (averaged) can be seen between scales 8 and 26, and the major modulating amplitude of average frequency 0.06 Hz between scales 72 and 222. The scalar energy is shown in Figure 5.8(c) and the times and scales of the uniform sections shown in Figure 5.8(d). Similarly, the signal for the infant at low risk for SIDS with the greatest amount of overall rhythmicity is shown in Figure 5.9. Again, the wavelet transform is shown only to 300 scales. Despite the very regular breathing of 0.5 Hz found between scales 6 and 28, amplitude modulation is absent, compared with the infant at high risk.

5.4.3 Discussion

The rhythmicity of the main breathing frequency and amplitude modulation were calculated, with uniform sections of the frequency components being successfully isolated and identified. The infants at low risk for SIDS showed very few uniform sections other than for the primary breathing frequency. This finding could have resulted from a low

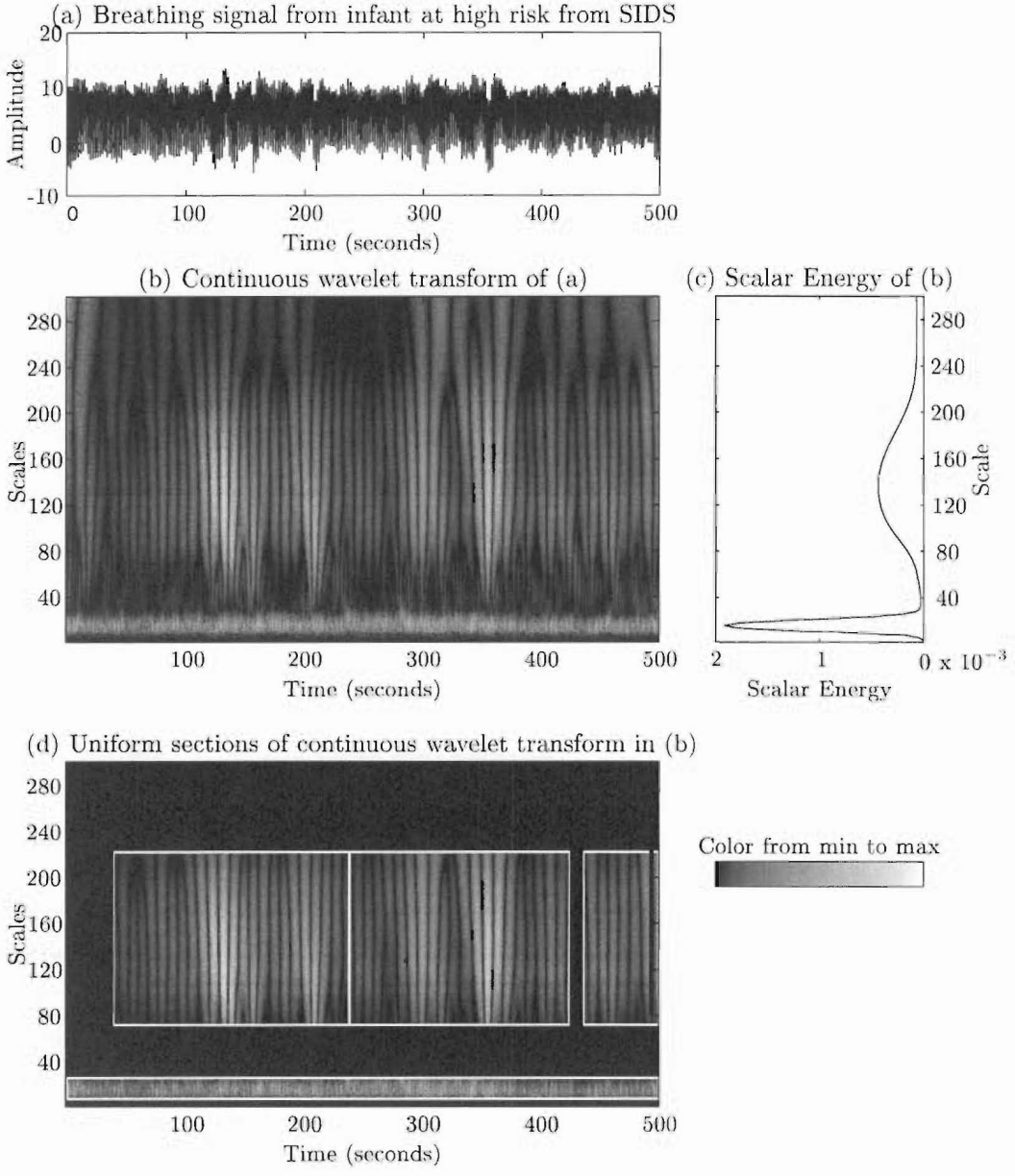


Figure 5.8 Infant at high risk for SIDS. (a) Breathing signal. Note the primary respiration component of the signal modulated by a slower frequency component. (b) Continuous wavelet transform of breathing signal that shows the two main frequencies. (c) Scalar energy of continuous wavelet transform. (d) Uniform sections of the continuous wavelet transform outlined in white. Scale range for the primary breathing frequency (0.5 Hz) is 8 to 26 and for the modulating frequency (0.06 Hz) is 72 to 222. Rhythmicity for the main breathing frequency is 26.6 % and for the modulating frequency is 57.2%.

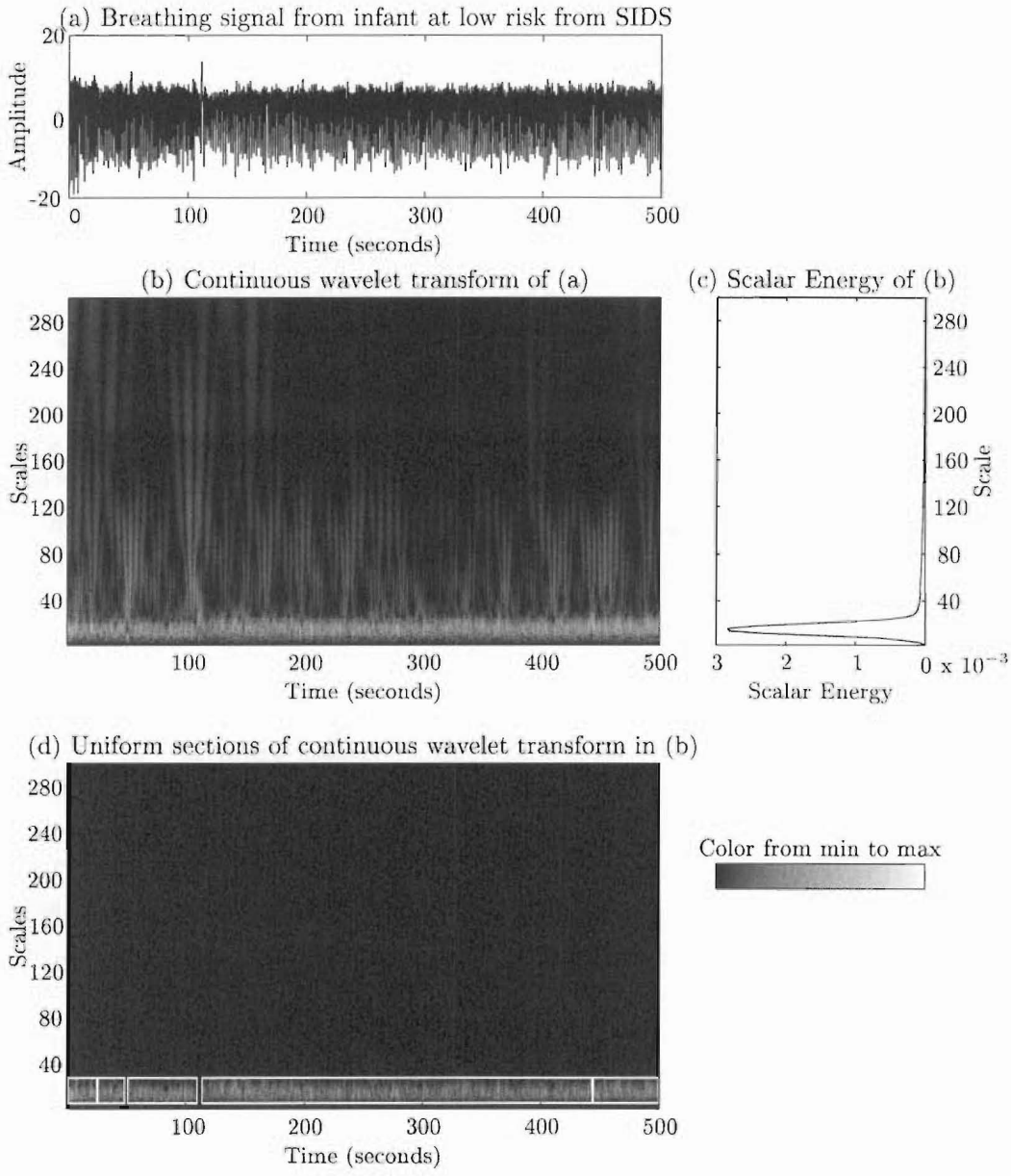


Figure 5.9 Infant at low risk for SIDS. (a) Breathing signal. Note the lack of modulation of the primary respiration. (b) Continuous wavelet transform of breathing signal. (c) Scalar energy of continuous wavelet transform. (d) Uniform sections of the continuous wavelet transform outlined in white. Scale range for the main breathing frequency (0.5 Hz) is 6 to 28. Rhythmicity for the main breathing frequency is 97.4 % with no modulating component.

magnitude amplitude modulation or from amplitude modulation which was not of a constant frequency. The isolation of the frequency components with constant frequency and sufficient magnitude ensured the relative importance of each component could be assessed. Having evaluated the relative importance of amplitude modulation in breathing, this evaluation was used as an objective measure of the differences between the high and low risk for SIDS groups.

The wavelet transform provides the basis for estimating rhythmicity. This provided an accurate evaluation of location and frequency of the various components of a signal. This is advantageous for two reasons: firstly, components lasting for only part of the signal are not evaluated as stationary components, but appropriately as transients; secondly, the timing of various frequency components may be important for understanding the mechanism of breathing control system operation in infants. One limitation with this method is that the calculation of scalar energy involves an averaging operation, which may result in the suppression of some transient components.

There is a wide range of rhythmicity in the primary frequency in both groups. This suggests that variance is high between individuals of both groups. Nevertheless, differences between high and low risk groups were apparent, suggesting that rhythmicity of amplitude modulation could be a useful measure for gaining insight into infant physiology.

Although the group is small, amplitude modulation occurred more in infants who had suffered an ALTE compared with controls. However, other factors such as age or time of night may also influence the results. Age may be a factor because the control of respiration is known to change as the infant matures [Haidmayer *et al.* 1982a]. Also, as infants mature, their sleep patterns during the night change. For instance, infants older than three months tend to have longer periods of REM sleep later in the night than their younger counterparts, indicating the respiratory control mechanisms in older infants have a circadian rhythm effect. This group was not large enough for statistical significance to test these hypotheses, but nevertheless, this initial study supports testing rhythmicity as a measure to differentiate between infants at high risk for SIDS and low risk for SIDS in a larger physiological study.

5.5 IMPACT OF AMPLITUDE MODULATION IN SIDS AND NORMAL INFANTS

This section presents the evaluation of the impact of amplitude modulation on the primary breathing frequency from infants who later succumbed to SIDS and their matched controls. The aim is to provide further results regarding the use of rhythmicity and impact to distinguish differences between the two groups. The relationship between age and amplitude modulation is also investigated as the respiratory control system of infants is known to mature during the first year of life [Haidmayer *et al.* 1982b]. Based

on the findings in Section 5.4, the infants who later succumbed from SIDS are expected to show a greater impact of amplitude modulation than their controls.

5.5.1 Methods

A set of data was recorded from five infants who later succumbed to SIDS and five control infants. The controls were matched for age, gender and birth weight. The infants were monitored for a 24 hour period. At the time of monitoring, the infants ranged from 9 days to 63 days of age. Respiration was measured at a sample rate of 15.6 Hz, and cardiac signals were monitored using thoracic wall leads [Richards *et al.* 1984]. The sleep state of each infant over the entire 24 hour period was classified, using the cardiac signals, either as awake, quiet sleep, Rapid Eye Movement (REM) sleep, or indeterminate state by means of an automatic state classification program [Harper *et al.* 1987].

For each infant, a 500 second epoch of quiet sleep was selected. Sections that encompassed periodic breathing, amplitude modulation, or a combination of both, were visually selected for analysis. The Daubechies symlet wavelet of order 8 was applied for scales 2 to 500 in steps of 2. Rhythmicity (Equation (5.1)) was calculated for the primary breathing frequency component and the amplitude modulation components. The impact (Equation (5.2)) of the amplitude modulation on the primary breathing frequency was calculated for each of the epochs.

A paired T-test was used to determine whether the impact of amplitude modulation on breathing was different for the two groups of infants.

5.5.2 Results

The impact of amplitude modulation on breathing (Equation (5.2)) for the ten infants is shown in Figure 5.10. The impact of the amplitude modulation was higher for the infants who later succumbed to SIDS for all pairs ($p < 0.01$ using a paired T-test). The small impact of amplitude modulation calculated for the control infants aged 24 and 40 days indicates either that the periodic breathing did not have uniform frequency, or had low energy in the wavelet domain relative to the energy of the primary breathing frequency.

An example of a breathing signal with periodic breathing from a control infant is shown in Figure 5.11(a), with the continuous wavelet transform shown in Figure 5.11(b). The periodic breathing is represented between scales 100 and 200. The scalar energy of the transform is shown in Figure 5.11(c). The primary breathing frequency is represented in the scalar energy as the peak at scale 14, and the amplitude modulation is represented in the second peak at scale 140. The uniform sections of the transform are

surrounded in white in Figure 5.11(d). The uniform sections do not completely span the time interval, illustrating the non-stationary aspect of the amplitude modulation.

5.5.3 Discussion

The results in this section indicate that normal infants have little uniform amplitude modulation of their breathing and suggests that infants who succumb to SIDS have more uniform amplitude modulation than their age matched pairs. There is a slight tendency for younger infants to exhibit a greater impact of amplitude modulation on the primary breathing frequency, although, with this small group, this change was not statistically significant. Physiologically, a change is expected, since the control systems of infants are known to change over this time [Haidmayer *et al.* 1982b].

These results are based on epochs of breathing in which periodic breathing or amplitude modulation was apparent. The small percentage of rhythmicity due to amplitude modulation in the normal infants suggests that either the frequency component did not have a significant amplitude, or the component did not have a constant frequency.

Infants who later succumb to SIDS show the same tendency as infants who have suffered an ALTE to exhibit greater uniform amplitude modulation than their controls. This finding suggests that deficiencies in breathing control may be similar in both SIDS and ALTE infants, an outcome that is supported by other studies linking ALTE and SIDS groups [Oren *et al.* 1986].

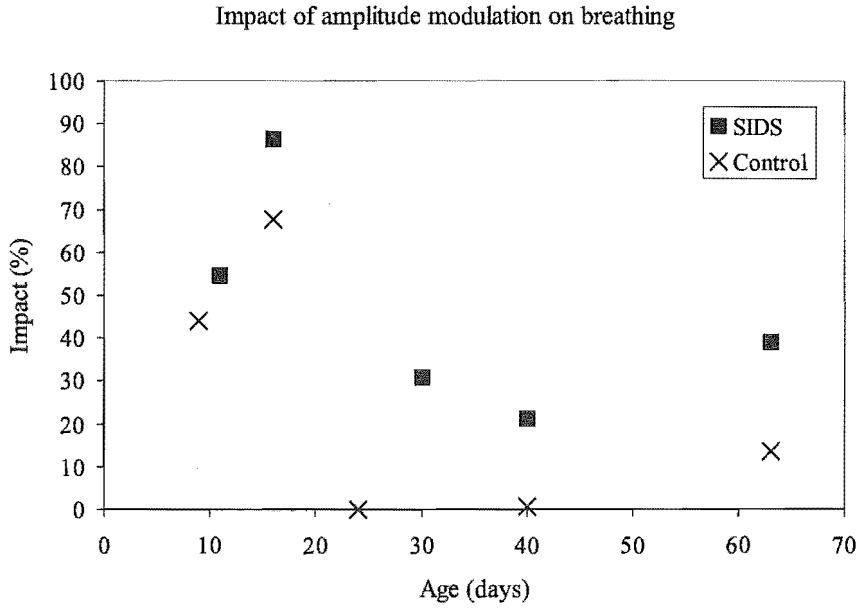


Figure 5.10 Impact of amplitude modulation on breathing, defined in Equation (5.2), for five infants who later succumbed to SIDS and five control infants.

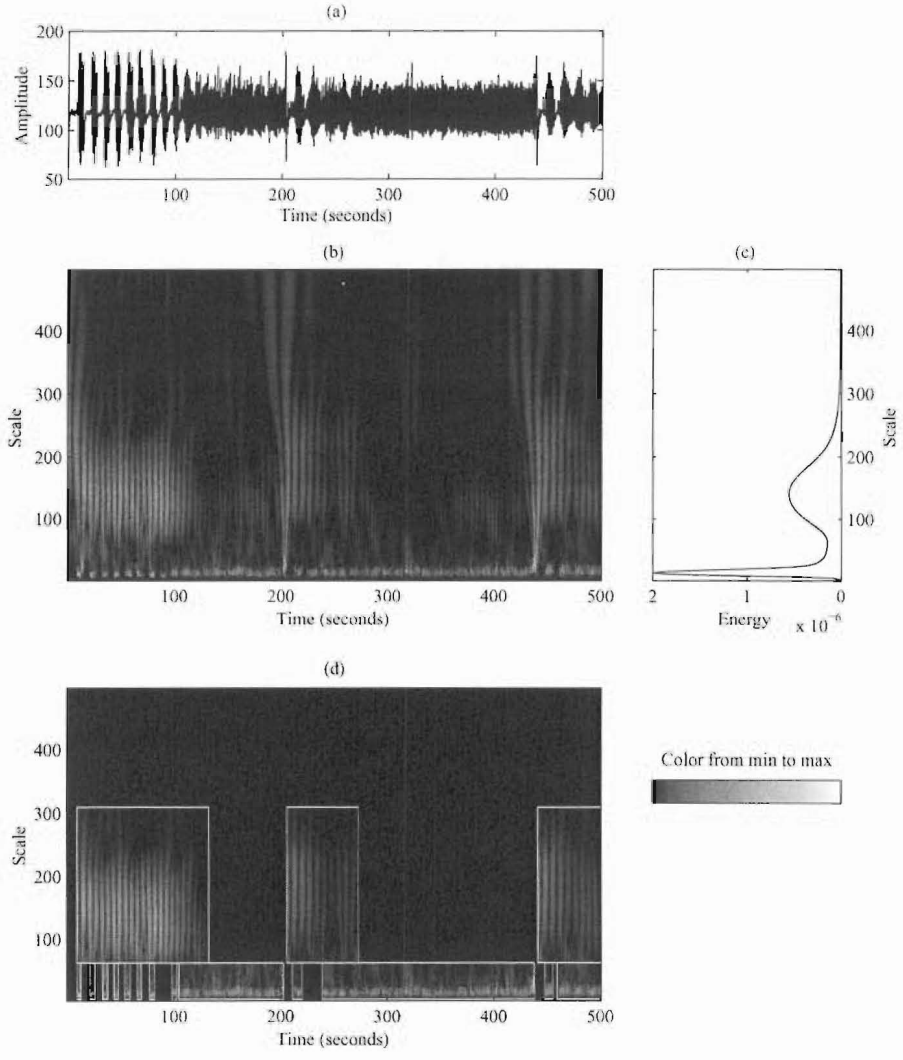


Figure 5.11 (a) Respiration epoch containing 3 episodes of periodic breathing from a control infant 16 days old. (b) Continuous wavelet transform of the signal. (c) Scalar energy of the continuous wavelet transform. (d) Uniform sections of the continuous wavelet transform outlined in white.

These results are consistent with spectral analysis of low (0.03 - 0.17 Hz) and high (0.3 - 1.3 Hz) frequency components of regular breathing from infants who later succumbed to SIDS and controls showing a higher ratio of low frequency to high frequency components in regular breathing in the SIDS infants [Rantonen *et al.* 1998]. An advantage of using rhythmicity is that the distinction between the primary breathing frequency and any amplitude modulation is based on the features of the scalar energy and is therefore determined by the data. This is in contrast to the spectral analysis done by Rantonen *et al.* [1998] in which the low and high frequency bands were predefined.

The physiological results point to a potentially fruitful area of research, in terms of both measurement of differences between groups and description of breathing control. Some caution must be applied to these findings, as this is a small sample, but the results do support the use of rhythmicity and impact as objective measures of amplitude modulation and the influence modulation has on respiratory signals of infants.

5.6 TIME OF NIGHT AND RHYTHMICITY IN SIDS AND CONTROL INFANTS

This section presents a more extensive study of infants who later succumbed to SIDS and their matched controls. Differences between the infants who later died and their controls are examined, and in particular, this study examines whether the rhythmicity of amplitude modulation is dependent on time of night. Time of night alters other physiological phenomena in infants, such as temperature [Brown *et al.* 1992] and heart rate variability [Schechtman and Harper 1991].

5.6.1 Methods

Thirteen infants who later succumbed to SIDS and 13 control infants, matched for age, gender and birth weight, were monitored for a 24 hour period at home. At the time of recording the infants ranged in age from 2 days to 63 days. Respiration was measured at a sample rate of 15.6 Hz, and cardiac signals were monitored using thoracic wall leads [Richards *et al.* 1984]. The sleep state of each infant over the entire 24 hour period was classified, using the cardiac signals, either as awake, quiet sleep, Rapid Eye Movement (REM) sleep, or indeterminate state by means of an automatic state classification program [Harper *et al.* 1987]. For each infant, two 500 second epochs of quiet sleep were selected. Each epoch occurred during different times of the night and were chosen, where possible, without movement artefacts, large sighs, apnoea, or periodic breathing.

Rhythmicity was calculated for both the primary breathing frequency and for amplitude modulation. The rhythmicity of the amplitude modulation was plotted against

the time of night of the start of the epoch. The rate of change of rhythmicity of the amplitude modulation over time was calculated.

A statistical comparison between the two groups was carried out for the rate of change of rhythmicity of the amplitude modulation over time using a Wilcoxon signed ranks test [Siegel 1956].

5.6.2 Results

The rhythmicity of amplitude modulation is shown against time of night for the control infants in Figure 5.12, and for the infants who succumbed to SIDS in Figure 5.13. The controls show greater rhythmicity of amplitude modulation early in the evening, and, in general, rhythmicity decreases during the night. The results for infants who succumbed to SIDS show a more even spread of rhythmicity throughout the night, but are more likely to increase rhythmicity of amplitude modulation later in the night.

The rate of change of rhythmicity over time between epochs is calculated for each infant and shown in Table 5.3. A positive rate of change represents more rhythmicity later in the night, and a negative rate of change represents less rhythmicity later in the night. The average rate of change for the control infants is -72.4 and the average rate of change for the SIDS infants is 12.2.

Comparing between groups, differences due to age or in rhythmicity of amplitude modulation of the regular, quiet sleep breathing were not statistically significant. However, infants who later succumbed to SIDS have a greater amount of amplitude modulation measured by rhythmicity as the time of night increases compared with the control infants ($p < 0.01$ by the Wilcoxon signed ranks test).

Control Infants		SIDS Infants	
Infant	Rate of change	Infant	Rate of change
1	-9.0	1	24.5
2	0.0	2	52.6
3	5.9	3	6.9
4	-92.8	4	-58.2
5	-476.2	5	85.0
6	0.0	6	19.9
7	-66.8	7	21.4
8	-95.9	8	-0.9
9	0.0	9	-8.2
10	-3.2	10	-31.6
11	-172.7	11	0.0
12	-65.9	12	1.5
13	34.9	13	45.8
Average	-72.4		12.2

Table 5.3 Slopes of rhythmicity vs time of night for control infants and infants who later succumbed to SIDS.

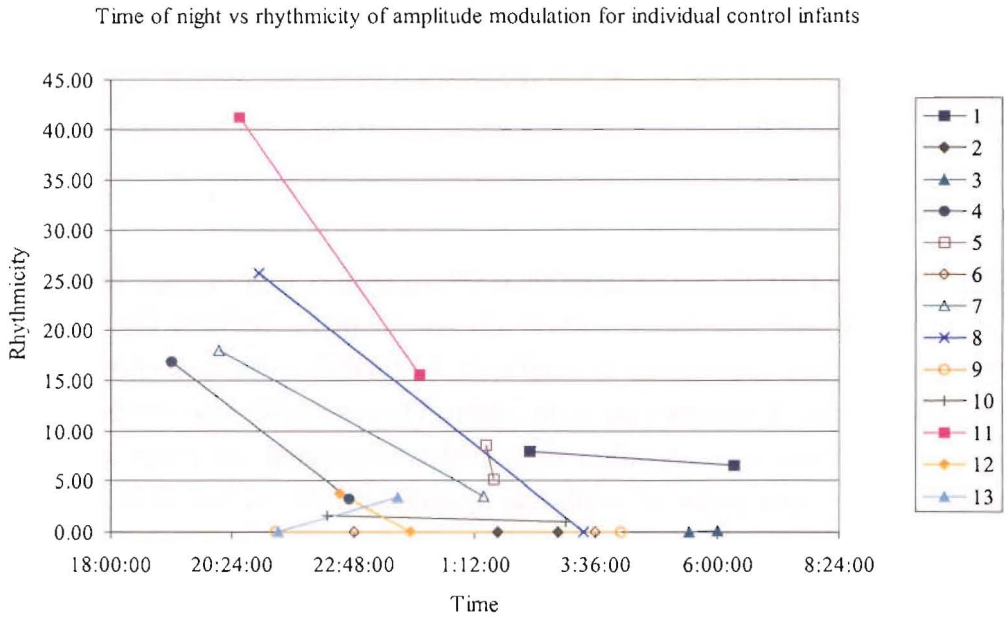


Figure 5.12 Rhythmicity of amplitude modulation of epoch of 500 s breathing from quiet sleep against time of night for 13 control infants.

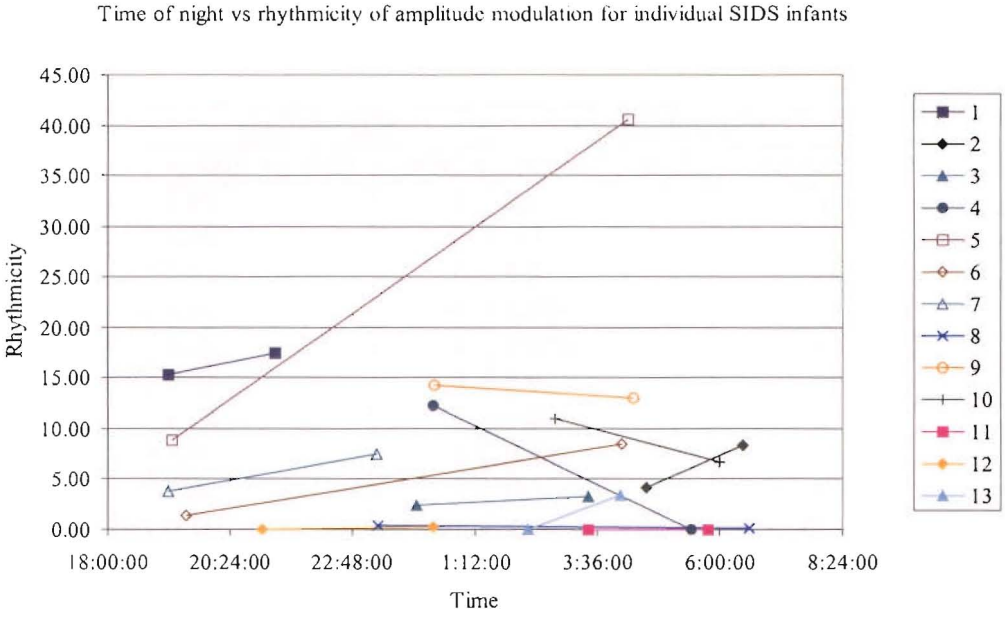


Figure 5.13 Rhythmicity of amplitude modulation of epoch of 500 s breathing from quiet sleep against time of night for 13 infants who later succumbed to SIDS.

5.6.3 Discussion

This quiet sleep breathing analysis investigated characteristics of breathing signals that were regular and, where possible, without sighs, apnoea or periodic breathing. Even so, some of these signals exhibited high rhythmicity of amplitude modulation frequency components.

Normal infants showed a tendency to exhibit less amplitude modulation later in the night compared with earlier in the night. Infants who later died, however, tended to show an increased amplitude modulation during the night. This suggests that infants who later succumb to SIDS have an abnormality of their breathing control, especially for slower frequency components.

Rantonen *et al.* [1998], who studied regular infant breathing by spectral analysis, found that infants who later succumbed to SIDS exhibited greater low frequency components relative to high frequency components in regular breathing than control infants. This finding is supported by the analysis in Section 5.5. However, in contrast to breathing which was selected, where possible, with obvious amplitude modulation, as in Sections 5.4 and 5.5, the analysis in this section showed no significant differences in general for rhythmicity between infants who later succumbed to SIDS and their matched controls. Instead, the finding of this study indicates that time of night is a confounding influence of rhythmicity. This finding may result from several potential mechanisms. One possibility is that the control mechanisms governing amplitude modulation perform differently for periodic breathing, or breathing showing obvious amplitude modulation, than when amplitude modulation is not obvious. It is also possible that, in general, the frequency of amplitude modulation in control infants normally has greater variability than that of infants who later succumb to SIDS, but that the variability of the amplitude modulation frequency is affected by the time of night. This would be similar to the primary breathing frequency, which shows greater variability in normal infants [Schechtman *et al.* 1988].

The dependence of rhythmicity on the time of night suggests a circadian rhythm influence. Several potential circadian factors may interact with modulation of breathing amplitude. Siblings of SIDS victims show an altered distribution of delta sleep across the night [Schechtman *et al.* 1994]; this deeper stage of sleep is suggestive of a reduced propensity to arouse. Infants who later succumb to SIDS indeed show a relative lack of micro arousals later in the night. Temperature in infants follows a circadian rhythm, with lower values at night. Increased core temperature in infants is related to periodic breathing and amplitude modulation; a substantial proportion of SIDS infants show higher than normal body temperatures at death [Stanton 1984]. The principal mechanism for cooling in humans is sweating, a sympathetic nervous system mediated behaviour. Two principal sympathetic nervous system mechanisms may potentially be operating in SIDS. The first, descending hypothalamic temperature drive, is known to

be affected by REM sleep [Ni *et al.* 1994]. Another system, however, is more suspect; this is the inferior olive/cerebellar/vestibular system which exerts substantial influence on sympathetic output. The inferior olive has consistently been shown to demonstrate dysplasia in SIDS victims [Storm *et al.* 1999], shows an absence of serotonergic receptors in SIDS victims [Panigrahy *et al.* In Press], and provides cerebellar projections that are very sensitive to hypoxic damage [Scholz 1963]. This system is thus involved in temperature control and deficiencies may be indicated by respiratory modulation.

5.7 CONCLUSIONS

Two main conclusions can be drawn from this chapter: firstly, rhythmicity based on the wavelet transform successfully measures amplitude modulation in an objective manner and secondly, amplitude modulation as measured by rhythmicity shows significant differences between infants who succumb from SIDS or are at high risk for SIDS and their controls.

Amplitude modulation is a non-stationary component of infant breathing signals during quiet sleep and may occur at several frequencies. The wavelet transform is shown to be a more suitable method for isolating these frequency components in the signal than the Fourier Transform or the Short Time Fourier Transform.

A measure based on the energy components of the wavelet transform, rhythmicity, is defined and shown to quantify the extent to which frequency components remain constant for the duration of the signal. Rhythmicity can be calculated for the signal as a whole, or for individual frequency components, such as the primary breathing frequency or frequencies resulting in amplitude modulation. A second measure, impact, is defined and is shown to quantify the influence frequency components have on the signal, measured in terms of the magnitude, duration and consistency of frequency.

The rhythmicity of amplitude modulation of the primary breathing frequency appears to be greater for both infants who had suffered an Apparent Life Threatening Event (ALTE) and infants who later succumbed to Sudden Infant Death Syndrome (SIDS) than controls in the respective preliminary studies. These results are based on epochs of breathing during quiet sleep in which periodic breathing or amplitude modulation were apparent in the signal. In the final study of infants who later succumbed to SIDS and their controls, epochs of breathing during quiet sleep in which periodic breathing or amplitude modulation were absent or minimised were analysed. In this study, the extent of amplitude modulation differed during the night. Control infants exhibited significant amplitude modulation, but only in the early portions of the night. Infants who later succumbed to SIDS, however, tended to increase the amount of amplitude modulation in their breathing during the night. These differences indicate the respiratory control systems of the two groups are also different.

In summary, by developing measures based on characteristics of signals in the wavelet domain, it is possible to accurately evaluate features of breathing signals that have previously only been approximated. These new measures may lead to greater understanding of disorders such as SIDS and to possible diagnosis of high or low risk for SIDS.

Chapter 6

CONSTRUCTION OF TWO-DIMENSIONAL WAVELETS

Two-dimensional wavelets are used to analyse images from a variety of sources [Mallat 1989, Daubechies 1992, Meyer 1993, Vetterli and Kovačević 1995]. Wavelet–based algorithms have been developed for processing medical images such as magnetic resonance images, x-ray mammograms and ultrasound scans of livers [Healy, Jr. and Weaver 1992, Healy, Jr. *et al.* 1995, Laine *et al.* 1995, Mojsilovič *et al.* 1997a and b]. Techniques that apply the wavelet transform have successfully been used for texture discrimination of ultrasound images [Mojsilovič *et al.* 1997a and b] and in enhancement of images [Healy, Jr. and Weaver 1992, Healy, Jr. *et al.* 1995, Laine *et al.* 1995]. Using wavelet transforms in the enhancement of images allows more flexibility of the enhancement process than traditional enhancement techniques such as Wiener filtering. A disadvantage of using wavelets is the increased computation required to process the enhancement function for all scales, but the rapid increase in affordable computing power over the last decade has encouraged the development of wavelet based enhancement.

In general, two-dimensional wavelets are classified as separable or non-separable. Separable wavelet transforms may be performed separately in each dimension, whereas non-separable wavelet transforms are performed simultaneously in all dimensions. Separable wavelets have the advantage of having lower computational complexity, but non-separable wavelets have more degrees of freedom in their design, potentially leading to wavelets more suited to the application [Kovačević and Vetterli 1995].

This chapter introduces two-dimensional wavelets and describes the theory for constructing separable and non-separable wavelets in the first and second sections, respectively. The third section describes image features in the wavelet domain for separable and non-separable wavelets and characteristics in the wavelet domain of several wavelets that will be used for image enhancement in Chapter 8. A summary of the chapter is given in the final section.

6.1 SEPARABLE WAVELETS

One defining characteristic of separable two-dimensional wavelets is that they may be constructed using one-dimensional wavelets [Mallat 1989, Rioul and Vetterli 1991, Daubechies 1992, Meyer 1993]. The output of the two-dimensional transform is separated into directional components that can provide better analysis of images with strong horizontal or vertical components such as images of buildings [Mallat 1989]. The separable wavelet transform leads to overall downsampling by 4 for each scale [Mallat 1989, Rioul and Vetterli 1991, Daubechies 1992].

6.1.1 Constructing Separable Wavelets

Separable wavelets in two-dimensions are constructed, following the multiresolutional approach described in Section 2.2.3, using one-dimensional spaces and their tensor products [Mallat 1989, Rioul and Vetterli 1991, Daubechies 1992, Meyer 1993]. Wavelets in higher dimensions may be constructed in a similar manner [Mallat 1989, Muraki 1993].

The procedure is as follows: choose a separable, finite energy, two-dimensional function that is a function of one-dimensional functions, $f(x)$ and $g(y)$ of each dimension x and y; i.e. $F(x,y) = f(x)g(y) \in L^2(\mathcal{R}^2)$. Let $\mathbf{V}_{2^j}$, where $j \in \mathcal{Z}$, be a subspace of $L^2(\mathcal{R}^2)$ that satisfies the two-dimensional case of properties (2.5) to (2.11). The two-dimensional space, $\mathbf{V}_{2^j}$, is defined as the tensor product, denoted by the symbol $\otimes$, of the one-dimensional space, V_{2^j} with itself:

$$\mathbf{V}_{2^j} = V_{2^j} \otimes V_{2^j}, \tag{6.1}$$

where V_{2^j} is a subspace of $L^2(\mathcal{R})$ satisfying the one-dimensional case of properties (2.5) to (2.11).

The approximation of a function $F(x,y)$ at a resolution 2^j is equal to its orthogonal projection on the vector space $\mathbf{V}_{2^j}$. There exists a unique scaling function, $\Phi(x,y)$, whose dilation and translation gives an orthonormal basis for $\mathbf{V}_{2^j}$. Similarly, there exists unique scaling functions, $\phi(x)$ and $\phi(y)$, whose dilations and translations each give an orthonormal basis for V_{2^j} [Mallat 1989, Daubechies 1992]. Using Equation (6.1), the two-dimensional scaling function can be written as:

$$\Phi(x,y) = \phi(x)\phi(y). \tag{6.2}$$

The complement space, $\mathbf{O}_{2^j}$, is defined to be the orthogonal complement of $\mathbf{V}_{2^j}$ in

$\mathbf{V}_{2^{j+1}}$. Therefore, $\mathbf{V}_{2^{j+1}}$ may be written as [Mallat 1989, Daubechies 1992]:

$$\begin{aligned}\mathbf{V}_{2^{j+1}} &= V_{2^{j+1}} \oplus V_{2^{j+1}} \\ &= (V_{2^j} \oplus O_{2^j}) \otimes (V_{2^j} \oplus O_{2^j}) \\ &= V_{2^j} \otimes V_{2^j} \oplus [(V_{2^j} \otimes O_{2^j}) \oplus (O_{2^j} \otimes V_{2^j}) \oplus (O_{2^j} \otimes O_{2^j})] \\ &= \mathbf{V}_{2^j} \oplus \mathbf{O}_{2^j}. \end{aligned} \tag{6.3}$$

where $\oplus$ denotes the appending of two vector spaces together.

It follows [Mallat 1989, Daubechies 1992] that an orthonormal basis of $\mathbf{O}_{2^j}$ can be defined using three wavelets given by:

$$\begin{aligned}\Psi_1(x,y) &= \phi(x)\psi(y), \\ \Psi_2(x,y) &= \psi(x)\phi(y), \\ \Psi_3(x,y) &= \psi(x)\psi(y). \end{aligned} \tag{6.4}$$

If the one-dimensional orthonormal bases give rise to compactly supported scaling functions and wavelets, then the two-dimensional orthonormal bases will also be compactly supported [Daubechies 1992]. Other properties of wavelets, such as orthogonality and biorthogonality, also transfer from one dimension to two [Daubechies 1992].

The sub-band filtering approach can be used to describe the two-dimensional separable wavelet transform. If an image is considered as an array, then the filtering functions are performed on the rows and columns of the array separately [Daubechies 1992], as illustrated in Figure 6.1. To iterate the process, the approximation, $a_{0,0}$, is taken as the input to the next level of filter banks, just as in the one-dimensional case (see Section 2.3.2).

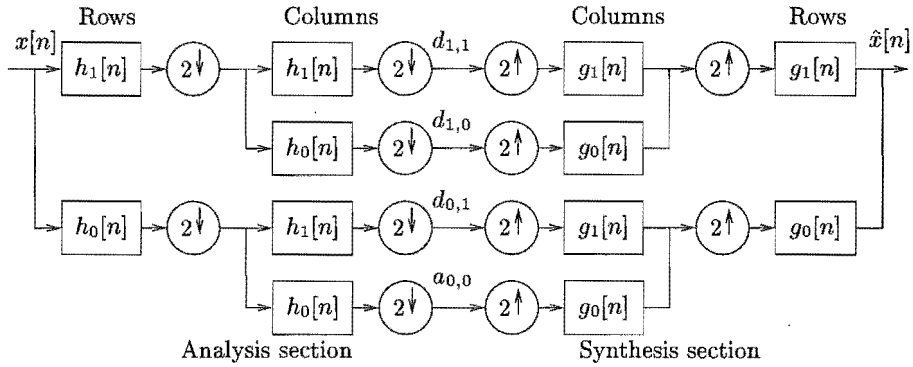


Figure 6.1 A two-dimensional separable analysis and synthesis filter bank.

The outputs of the analysis section of the filter bank are two-dimensional images. For a given level, the output image size is one quarter of the size of the input to that level, because downsampling by 2 occurs in both dimensions, leading to overall downsampling by 4. As in the one-dimensional discrete wavelet transform case, it is

customary to display the details at each level and the approximation at the final level. A typical organisation for this display is shown in Figure 6.2.

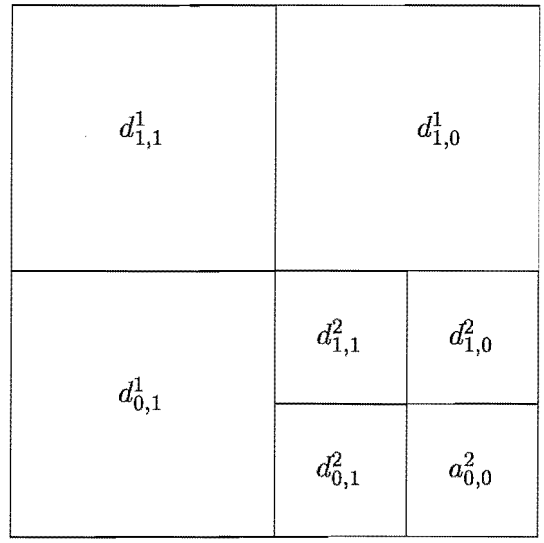


Figure 6.2 A typical organisation for displaying the two-dimensional discrete wavelet transform.

6.2 NONSEPARABLE WAVELETS

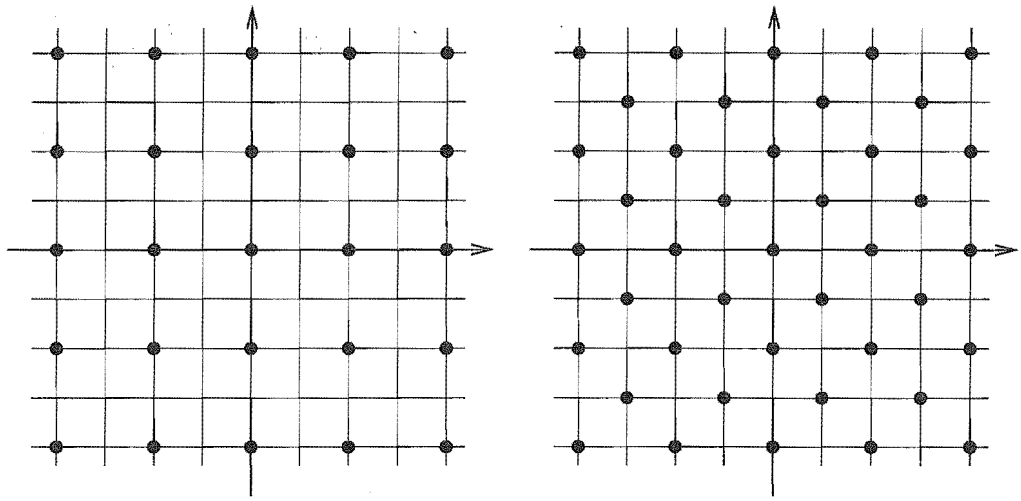
Non-separable wavelets have several characteristics that distinguish them from separable wavelets, one of which is that a non-separable wavelet cannot be constructed using one-dimensional wavelets. Numeric expressions for non-separable wavelets are more complex to derive than for separable wavelets. However, non-separable wavelets give greater freedom in designing sub-sampling schemes [Kovačević and Vetterli 1995]. One such scheme described in this section is quincunx sub-sampling, which leads to an overall downsampling by 2. Downsampling by 2, compared with downsampling by 4 in the separable case, results in smaller bandwidths in the frequency domain for each level giving greater control in enhancement schemes. It is also possible to design non-separable filters that are non-directional. Using non-directional wavelet transforms is more appropriate for images of natural objects, such as medical images, as these images often do not have inherent directional biases. These characteristics are explained in more detail in the following sections.

6.2.1 Lattices

A lattice describes how a signal is sampled. The first time a signal is sampled, the lattice is known as the *initial* lattice, with subsequent sampling, termed *sub-sampling*, performed on *sub-lattices*. Figure 6.3(a) shows an initial lattice with sampling points represented by the crossing of horizontal and vertical lines, or index points. The two-

dimensional signals considered in this thesis are initially sampled using a rectangular lattice.

When performing two-dimensional wavelet transforms, the signal is filtered and then downsampled. The downsampling can be described using a sub-lattice of the initial lattice [Kovačević and Vetterli 1992]. The sub-lattice resulting from separable two-dimensional sub-sampling is represented by solid dots in Figure 6.3(a). There is an overall downsampling rate of 4, i.e., only one in every four samples is kept. If, as in the one-dimensional case, a downsampling rate of 2 is desired, then a different sub-lattice is required. One solution is the quincunx sub-lattice [Rioul and Vetterli 1991, Kovačević and Vetterli 1992] illustrated in Figure 6.3(b). The quincunx sub-lattice gives an overall sub-sampling rate of 2 whilst still reducing the number of samples in both the horizontal and vertical directions. It is a non-separable sub-lattice [Kovačević and Vetterli 1992, Kovačević and Vetterli 1995].



(a) Original sampling lattice represented by crossing of horizontal and vertical lines. A separable sub-lattice is represented by solid dots.

(b) Original sampling lattice represented by crossing of horizontal and vertical lines. A non-separable quincunx sub-lattice is represented by solid dots.

Figure 6.3 Initial sampling and sub-sampling lattices; (a) separable sub-sampling, and (b) quincunx sub-sampling.

Sub-lattices can be represented by a dilation matrix $\mathbf{D}$ [Kovačević and Vetterli 1992, Kovačević and Vetterli 1995]. The index points of the sub-lattice are given by weighted integer combinations of the columns of $\mathbf{D}$. A separable lattice is a lattice that can be represented by a diagonal matrix. For example, the separable sub-lattice in Figure 6.3(a) may be represented by $\mathbf{D}_0 = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right)$. The sampling rate, N, is given by $N = |Det(\mathbf{D}_i)|$. Notice that $N_0 = |Det(\mathbf{D}_0)| = 4$.

The matrix $\mathbf{D}$ is not unique for any given lattice. Some examples of dilation

matrices for the non-separable quincunx lattice are:

$$\mathbf{D}_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{D}_2 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{D}_3 = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}. \tag{6.5}$$

A dilation matrix in two-dimensional downsampling takes the place of the dilation function in the one-dimensional two-scale difference equation for wavelets, Equation (2.29), leading to the two-dimensional two-scale difference equation:

$$\Psi_i(\mathbf{t}) = \sqrt{N} \sum_{k=\mathcal{Z}^n} h_i[k]\phi(\mathbf{Dt} - \mathbf{k}). \tag{6.6}$$

When iterating the filter bank, sub-sampling by $\mathbf{D}$ is also iterated. Therefore, different dilation matrices can produce very different results [Kovačević and Vetterli 1992]. For instance, iterating $\mathbf{D}_1$ twice results in a diagonal matrix equivalent to iterating $\mathbf{D}_0$ once and, therefore, results in separable sub-sampling every second iteration. Iterating $\mathbf{D}_3$, however, will never result in a diagonal matrix and, therefore, will never result in separable sub-sampling.

Another consideration in choosing the matrix $\mathbf{D}$ is whether it represents a dilation in all directions. If it does, the associated wavelet will also increase its size in all dimensions over successive iterations. Dilation in all dimensions is equivalent to requiring that all eigenvalues of $\mathbf{D}$ should have a magnitude strictly greater than one [Kovačević and Vetterli 1992].

In this thesis, the dilation matrix $\mathbf{D}_1$ is used as it gives equal sub-sampling in both dimensions, and every second iteration is equivalent to separable sub-sampling, allowing comparison between the two classes of two-dimensional wavelets.

The polyphase domain, as explained in Section 3.1, will be used extensively in the remainder of this section to explain the construction of two-dimensional wavelets. The polyphase matrix, $H_p(\mathbf{z})$, is given by $H_p(\mathbf{z}^{\mathbf{D}})$ in the upsampled domain, where $\mathbf{z}^{\mathbf{D}}$ is given by

$$\mathbf{z}^{\mathbf{D}} = (z_1, z_2)^{\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)} = \left(z_1 z_2, z_1 z_2^{-1}\right). \tag{6.7}$$

6.2.2 Constructing Non-Separable Wavelets

Two-dimensional non-separable wavelets are often constructed using cascade structures, which allow high order filters to be derived from lower order filters [Vetterli and Kovačević 1995]. Two advantages of using cascade structures to construct filters are, firstly, that the filters usually have low complexity, and secondly, that the filter coefficients can be quantised without affecting the desired form [Vetterli and Kovačević 1995]. One disadvantage is that very few results on completeness of cascade

structures in multi-dimensions exist, unlike in the one-dimensional case.

Cascades of orthogonal building blocks produce orthogonal filter banks [Kovačević and Vetterli 1992]. The building blocks consist of orthogonal matrices and diagonal delay matrices. Linear phase filter banks may be constructed using linear phase building blocks.

Different sampling matrices for the same lattice will greatly affect the geometry of the filters obtained [Vetterli and Kovačević 1995]. Using $\mathbf{D}_1$, the following cascade can be used to generate either orthogonal or linear phase filters based on the quincunx sampling lattice:

$$H_p(z_1, z_2) = \left[\prod_{i=K-1}^{1} R_{2_i} \begin{pmatrix} 1 & 0 \\ 0 & z_2^{-1} \end{pmatrix} R_{1_i} \begin{pmatrix} 1 & 0 \\ 0 & z_1^{-1} \end{pmatrix} \right] R_0. \tag{6.8}$$

To obtain orthogonal filters, the R_{j_i} must be unitary. A matrix, A, is unitary if the inverse of the matrix is equal to the conjugate transpose of the matrix, that is, $A^{-1} = (A^*)^T$, where A^* has entries that are the complex conjugates of the corresponding entries in A [Anton 1984].

To obtain linear phase filters, R_{j_i} must be symmetric. In order to check that a filter is linear phase, a testing condition is used. A testing condition for a linear phase filter can be given either in the spatial domain or in the polyphase domain [Karlsson and Vetterli 1990]. To test in the spatial domain, assume there is a filter, m, specified on the input lattice, with the two-dimensional impulse response coefficients given by $h_m[n_1, n_2]$. Then, if this filter has a finite size of $N_1 \text{x} N_2$, and is linear phase, the filter coefficients will have the following symmetry:

$$h_m[n_1, n_2] = \pm h_m[N_1 - n_1 - 1, N_2 - n_2 - 1], \tag{6.9}$$

where the sign change occurs when the impulse response has odd symmetry. The sign is the same for all the coefficients of a given filter.

In the polyphase domain, the linear phase testing condition may be written as:

$$H_m(z_1, z_2) = \pm z_1^{-N_1+1} z_2^{-N_2+1} H_m(z_1^{-1}, z_2^{-1}). \tag{6.10}$$

The desired degree of regularity is guaranteed in the one-dimensional case by imposing a zero of a sufficiently high order at π in the Fourier Transform of the filter. This is described for the one-dimensional case in Section 3.3 using polyphase notation and is equivalent to setting $H_0(z) = 0$ for $z = -1$. In the two-dimensional case, imposing a zero of the highest possible order at (π, π) on the Fourier Transform of the low pass filter, $H_0(\omega_1, \omega_2)$, is possible only for small filters, as the solution is otherwise non-trivial [Vetterli and Kovačević 1995]. Imposing a zero at (π, π) on the two-dimensional

low pass filter is given by

$$\left.\frac{\partial^{k-1}H_0(\omega_1,\omega_2)}{\partial^l\omega_1\partial^{k-l-1}\omega_2}\right|_{(\pi,\pi)}=0,\quad\begin{cases}k=1,...,m,\\ l=0,...,k-1.\end{cases}\tag{6.11}$$

6.2.2.1 Orthogonal Wavelets

In this section, the construction of orthogonal two-dimensional non-separable wavelets is described using the cascade in Equation (6.8). An example and a solution for the smallest regular two-dimensional filter is given.

For orthogonal two-dimensional non-separable filters, the cascade in Equation (6.8) is used with R_{j_i} chosen to be unitary. One choice for R_{j_i} is

$$R_{j_i}=\begin{pmatrix}1 & -a_j\\ a_j & 1\end{pmatrix}.\tag{6.12}$$

In the polyphase domain, the filter coefficients are given by

$$H(\mathbf{z})=\begin{pmatrix}H_0(\mathbf{z})\\ H_1(\mathbf{z})\end{pmatrix}=H_p(\mathbf{z}^{\mathbf{D}})\begin{pmatrix}1\\ z_1^{-1}\end{pmatrix}.\tag{6.13}$$

where $H_p(\mathbf{z})$ is given by Equation (6.8) and $\mathbf{z}^{\mathbf{D}}$ is given by Equation (6.7). The filter coefficients, $h_0[n_1,n_2]$ and $h_1[n_1,n_2]$ for the analysis low pass and high pass filters respectively, are given in the spatial domain using Equation (3.1):

$$\begin{aligned}h_0[n_1,n_2]&=\begin{pmatrix}0 & -a_1 & -a_0a_1 & 0\\ -a_2 & -a_0a_2 & -a_0 & 1\\ 0 & a_0a_1a_2 & -a_1a_2 & 0\end{pmatrix},\\ h_1[n_1,n_2]&=c\begin{pmatrix}0 & -a_1a_2 & -a_0a_1a_2 & 0\\ 1 & a_0 & -a_0a_2 & a_2\\ 0 & -a_0a_1 & a_1 & 0\end{pmatrix},\end{aligned}\tag{6.14}$$

where $c=\frac{1}{\sqrt{a_0a_1a_2}}$ and is a normalising factor to ensure perfect reconstruction. Using Equation (6.11) and imposing a second order zero, the following solutions are obtained:

$$a_0=\pm\sqrt{3},\quad a_1=0,\qquad a_2=2\pm\sqrt{3},\tag{6.15}$$

$$a_0=\mp\sqrt{3},\quad a_1=\mp\sqrt{3},\qquad a_2=2\pm\sqrt{3}.\tag{6.16}$$

These solutions should be scaled by a factor of $\frac{(1-\sqrt{3})}{4\sqrt{2}}$ for normalisation. The solution obtained in Equation (6.15) is the one-dimensional Daubechies daub2 filter. The solution obtained in Equation (6.16) is a counterpart to daub2. It has been shown that this second solution is the smallest regular two-dimensional filter [Villemoes 1994,

Vetterli and Kovačević 1995] resulting in continuous scaling and wavelet functions.

The synthesis filters are obtained by modulation and time reversal, as in Equations (3.21), giving:

$$g_0[n_1,n_2] = \begin{pmatrix} 0 & -a_1a_2 & a_0a_1a_2 & 0 \\ 1 & -a_0 & -a_0a_2 & -a_2 \\ 0 & -a_0a_1 & -a_1 & 0 \end{pmatrix},$$

$$g_1[n_1,n_2] = \begin{pmatrix} 0 & a_1 & -a_0a_1 & 0 \\ a_2 & -a_0a_2 & a_0 & 1 \\ 0 & -a_0a_1a_2 & -a_1a_2 & 0 \end{pmatrix}. \tag{6.17}$$

These filters will be denoted *kovaorth* and will be used in Section 6.3 and Chapter 8.

6.2.2.2 Linear Phase Wavelets

In this section, linear phase two-dimensional non-separable filters are constructed. The filters described here are used in Section 6.3 and Chapter 8.

Filters are often referred to by the shape made by their non-zero coefficients. Linear phase filters that are nearly a diamond shape, and of equal sizes, can be constructed using the cascade in Equation (6.8), with $R_{j_i} = \begin{pmatrix} 1 & a_j \\ a_j & 1 \end{pmatrix}$. The following linear phase filters result:

$$h_0[n_1,n_2] = \begin{pmatrix} 0 & a_1 & a_0a_1 & 0 \\ 1 & a_0 & a_0 & 1 \\ 0 & a_0a_1 & a_1 & 0 \end{pmatrix},$$

$$h_1[n_1,n_2] = \begin{pmatrix} 0 & -a_1 & -a_0a_1 & 0 \\ 1 & a_0 & -a_0 & 1 \\ 0 & a_0a_1 & a_1 & 0 \end{pmatrix}. \tag{6.18}$$

These filters have $a_2 = 1$, which means that they are symmetric. Note from Equation (6.10) that the filters have centro-symmetry, that is they can be rotated 180° about their centre with no effect, other than a possible multiplication by -1. However, the filters do not have top-bottom or left-right symmetry.

Diamond shape filters with top-bottom and right-left symmetry produce a transform that has undistorted locational information in the wavelet domain. If diamond shape filters with top-bottom and right-left symmetry are desired, then it is necessary to have different sized filters.

To construct multidimensional filters with a perfect diamond shape, an iterative process is used on the linear phase filter pairs developed in Section 3.4. The filters have a length of $2K+1$ and $2K+3$ and are the same size in all dimensions. In the following equations, $\mathbf{z} = (z_1, \ldots z_n)$ in n dimensions is denoted by $\mathbf{z}_{(n)}$, and $\mathbf{z}$ in the upsampled

domain is denoted by $\mathbf{z}^{\mathbf{D}}$ as in Equation (6.11). The general solution in n-dimensions of the filter pair described by Equation (3.51) is given by [Kovačević and Vetterli 1992]:

$$\begin{aligned}
H_{00}(\mathbf{z}^{\mathbf{D}}_{(n)}) &= H_{00}(\mathbf{z}^{\mathbf{D}}_{(n-1)}) + a\, z_1^{-1}(z_n^{-1} + z_n^1), \\
H_{01}(\mathbf{z}^{\mathbf{D}}_{(n)}) &= H_{01}(\mathbf{z}^{\mathbf{D}}_{(n-1)}), \\
H_{10}(\mathbf{z}^{\mathbf{D}}_{(n)}) &= H_{10}(\mathbf{z}^{\mathbf{D}}_{(n-1)}) + H_c(\mathbf{z}^{\mathbf{D}}_{(n-1)})\, z_1^{-1}(z_n^{-1} + z_n^1) + c\, z_1^{-2}(z_n^{-2} + z_n^2), \\
H_{11}(\mathbf{z}^{\mathbf{D}}_{(n)}) &= H_{11}(\mathbf{z}^{\mathbf{D}}_{(n-1)}) + d\, z_1^{-1}(z_n^{-1} + z_n^1),
\end{aligned} \tag{6.19}$$

where

$$\begin{aligned}
H_{01}(\mathbf{z}^{\mathbf{D}}_{(n-1)}) &= b,\ b \neq 0, \\
H_c(\mathbf{z}^{\mathbf{D}}_{(n-1)}) &= \frac{d\, H_{00}(\mathbf{z}^{\mathbf{D}}_{(n-1)}) + a\, H_{11}(\mathbf{z}^{\mathbf{D}}_{(n-1)})}{b}, \\
a\, d &= b\, c, \\
t_{n-1} + 2ad &\neq 0,
\end{aligned}$$

and where $Det[H_p(\mathbf{z}^{\mathbf{D}}_{(n-1)})] = t_{n-1}\, z_1^{-2}$. The only non-zero coefficient in the determinant of the polyphase matrix in $n-1$ dimensions is t_{n-1}. The determinant of the polyphase matrix in n dimensions is $Det[H_p(\mathbf{z}^{\mathbf{D}}_{(n)})] = (t_{n-1} + 2ad)\, z_1^{-2} = t_n\, z_1^{-2}$.

Using Equation (3.51) in the upsampled domain, i.e. $z_1 \Rightarrow z_1^2$, and Equation (6.19), a solution for $n = 2$ can be found. First an expression for H_c is evaluated:

$$\begin{aligned}
H_c(z_1^2) &= \frac{d\,(1 + z_1^{-2}) + a\, a_1(1 + z_1^{-2})}{a_1}, \\
&= \left(a + \frac{d}{a_1}\right)(1 + z_1^{-2}),
\end{aligned} \tag{6.20}$$

and second, c is expressed as a function of the other variables, $c = \frac{ad}{a_1}$, leading to the solution for the polyphase components in the upsampled domain for 2 dimensions [Kovačević and Vetterli 1992]:

$$\begin{aligned}
H_{00}(\mathbf{z}^{\mathbf{D}}) &= 1 + z_1^{-2} + a\, z_1^{-1}(z_2^{-1} + z_2^1), \\
H_{01}(\mathbf{z}^{\mathbf{D}}) &= a_1, \\
H_{10}(\mathbf{z}^{\mathbf{D}}) &= 1 + a_2 z_1^{-2} + z_1^{-4} + \left(a + \frac{d}{a_1}\right)(1 + z_1^{-2}) z_1^{-1}(z_2^{-1} + z_2^1) + \frac{d}{a_1}\, z_1^{-2}(z_2^{-2} + z_2^2), \\
H_{11}(\mathbf{z}^{\mathbf{D}}) &= a_1(1 + z_1^{-2}) + d\, z_1^{-1}(z_2^{-1} + z_2^1).
\end{aligned} \tag{6.21}$$

This solution leads to the following impulse response filters using Equation (3.1):

$$h_0[n_1,n_2] = \begin{pmatrix} & a & \\ 1 & a_1 & 1 \\ & a & \end{pmatrix},$$

$$h_1[n_1,n_2] = \begin{pmatrix} & & \frac{ad}{a_1} & & \\ & a+\frac{d}{a_1} & d & a+\frac{d}{a_1} & \\ 1 & a_1 & a_2 & a_1 & 1 \\ & a+\frac{d}{a_1} & d & a+\frac{d}{a_1} & \\ & & \frac{ad}{a_1} & & \end{pmatrix}. \tag{6.22}$$

Imposing circular symmetry and a second order zero at (π,π) using Equation (6.11) gives the results:

$$h_0[n_1,n_2] = \begin{pmatrix} & 1 & \\ 1 & 4 & 1 \\ & 1 & \end{pmatrix},$$

$$h_1[n_1,n_2] = \begin{pmatrix} & & 1 & & \\ & 2 & 4 & 2 & \\ 1 & 4 & -28 & 4 & 1 \\ & 2 & 4 & 2 & \\ & & 1 & & \end{pmatrix}. \tag{6.23}$$

The synthesis filters are found using Equation (3.18):

$$g_0[n_1,n_2] = \begin{pmatrix} & & 1 & & \\ & 2 & -4 & 2 & \\ 1 & -4 & -28 & -4 & 1 \\ & 2 & -4 & 2 & \\ & & 1 & & \end{pmatrix},$$

$$g_1[n_1,n_2] = \begin{pmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{pmatrix}. \tag{6.24}$$

Unfortunately, this choice of synthesis low pass filter leads to a discontinuous function [Kovačević and Vetterli 1992]. However, it is possible to swap the analysis and synthesis filters using Equation (3.11), leading to a very regular low pass synthesis filter $g_0' = \begin{pmatrix} & 1 & \\ 1 & 4 & 1 \\ & 1 & \end{pmatrix}$. Having a regular synthesis filter is desirable as the output is perceptually better [Misiti *et al.* 1996]. The filters g_0' and corresponding g_1' will be called *kovalinp* are used in the next section and in Chapter 8 to analyse and enhance images.

6.3 ASPECTS OF TWO-DIMENSIONAL WAVELET TRANSFORMS

In this section, three types of images are examined: the first is a fabricated test image made of ellipses; the second image is from a photograph of the rose window in Christchurch Cathedral; and the third is an ultrasound image. These images are transformed using separable and non-separable wavelets and shown in the wavelet domain to illustrate pertinent features of two-dimensional analysis. Wavelet domain characteristics of four wavelets are presented.

6.3.1 Features of Images in the Wavelet Domain

In this section, examples of transforming images into the wavelet domain using both linear phase separable and non-separable wavelets are presented. The separable examples follow the convention set out in Figure 6.2 for displaying separable two-dimensional wavelet transforms. The first image to be transformed is a simple one, allowing the reader to identify pertinent features in the wavelet domain. The second test image contains more details. The wavelet transforms of this image using both separable and non-separable wavelets are displayed, allowing the reader to examine a more complex transform. The third example is an ultrasound image. The overall aim is to identify features from the original images as they appear in the wavelet transform, and the scales at which the features occur.

Linear phase wavelets are used to preserve the position of features in the wavelet domain relative to the original image. The separable wavelet is constructed using the tensor product of the bior1.3 one-dimensional wavelet in both directions, as in Section 6.1. The non-separable wavelet used is kovalinp, the diamond-shaped filter given in Section 6.2.2.2. The bior1.3 wavelet was chosen so the two-dimensional wavelet formed from the tensor product would be approximately the same size as the kovalinp wavelet. These two wavelets are used in Chapter 8. The edge effects of the transform are negated by using symmetrical extension of the images prior to transformation [Chen Li *et al.* 1995]. The symmetrical extension is visible in all of the transforms.

The first test image consists of four ellipses offset in both the x- and y-axes and is used to illustrate the transformation of a simple figure into the wavelet domain. The original 256x256 image is shown in Figure 6.4. The image ranges in intensity, I, from black, $I = 0$, to white, $I = 1$. The background has intensity $I = 0.5$. The top, left-hand ellipse, ellipse one, has intensity $I = 1$. The top, right-hand ellipse, ellipse two, has intensity $I = 0$. The bottom, left-hand ellipse, ellipse three, has intensity $I = 0.75$, and contains a very small area of intensity $I = 1$. The bottom right-hand ellipse, ellipse four, has intensity $I = 0$, and also contains a small area of intensity $I = 1$. The features to look for in the following transforms are the edges of the ellipses and the small high

intensity areas. As a linear phase wavelet is being used, the features of interest appear in the same relative location within the individual parts of the transform as in the original image.

The result of transforming the image into the wavelet domain using a separable wavelet is shown in Figure 6.5. Each detail component of the transform has been normalised so that the features within each component can be seen. The approximation $a_{0,0}$ has a much higher magnitude than the detail parts. Three levels of decomposition are shown. The detail $d^j_{1,1}$ results are on the leading diagonal, highlighting the diagonal components of the edges of the ellipses. The detail $d^j_{0,1}$ results are below the corresponding $d^j_{1,1}$ results, highlighting vertical components of the edges of the ellipses of the image. The $d^j_{1,0}$ results are to the left of the $d^j_{1,1}$ results and they highlight the horizontal components of the edges of the ellipses. The $a_{0,0}$ result, the approximation, is in the bottom right hand corner. The small areas of high intensity in ellipses three and four can only be seen in the $d^j_{0,1}$ results. The area in ellipse three can only just be made out in scale $j = 1$, whereas the area in ellipse four can be made out at all scales, including in the approximation. Large features, such as the area in ellipse four, are prominent in the approximation.

The results of transforming the ellipse image into the wavelet domain using a non-separable wavelet are shown in Figure 6.6. Six levels of decomposition are shown. The sixth level of decomposition has a resolution that is the same as the third level of decomposition in the separable case. The decompositions are shown at the same size as the original. Note that the resolution decreases as the scale parameter increases. The transforms with an odd valued scale parameter are shown rotated by 45°; rotating the transform of odd scales by 45° ensures the transform is on a rectangular grid, which is easily imaged. As with the separable wavelet transform, the edges of the ellipses are the notable features. However, each part of the transform shows the edges of the ellipses in all orientations, without directional preferences, which would be an advantage for detecting edges in the wavelet domain. The area of high intensity in ellipse three is just distinguishable in scale $j = 2$ (Figure 6.6(c)). The area of high intensity in ellipse four is clearly distinguishable in scales $j = 1$ to $j = 6$. Very small objects are hard to detect next to edges, especially if they have a similar intensity to their surrounding area. It is useful to identify features in isolated scales in order to be able to enhance them. There is more opportunity for isolating certain features in individual scales in the non-separable wavelet domain, as the change in resolution between decomposition levels is half that of separable wavelets.

Several real life images were considered for the second test image, but one was chosen because of well defined lines in all orientations. The chosen image was scanned from a photograph of the rose window in the west wall in Christchurch Cathedral and is shown in Figure 6.7. The image was saved at a resolution of 256x256 pixels.

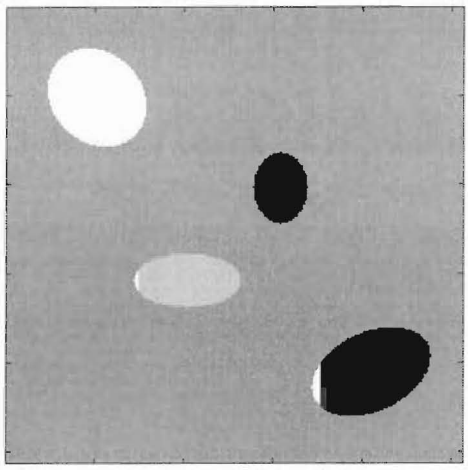


Figure 6.4 Original ellipse test image. The ellipses are numbered, starting from the top of the image, from one to four.

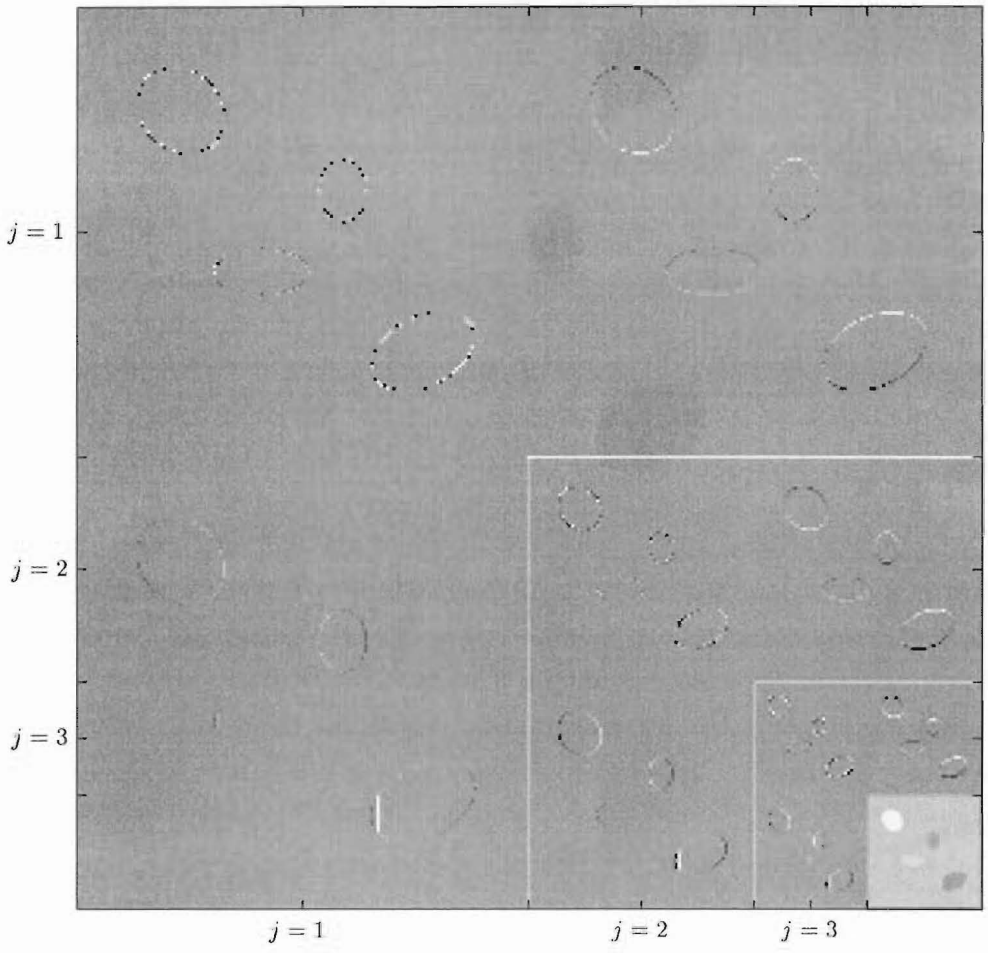
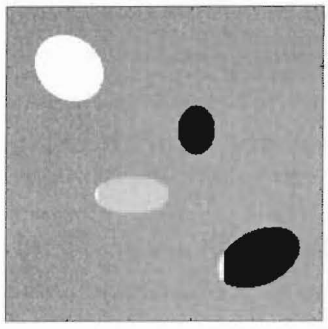
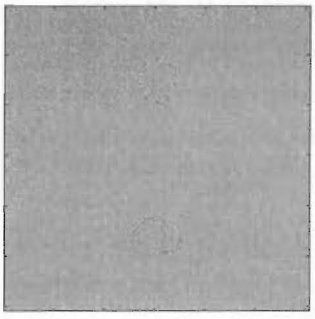


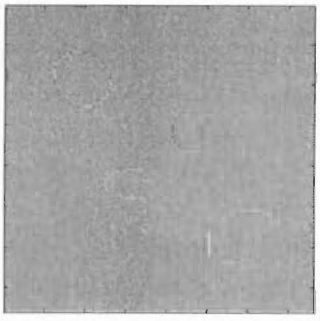
Figure 6.5 Separable wavelet transform of the ellipse test image, showing three scales, j, of decomposition.



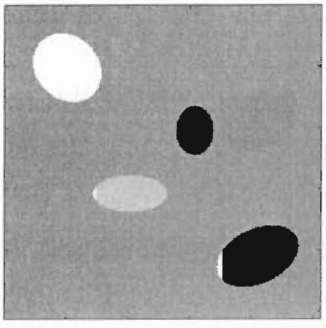
(a) Original ellipse test image.



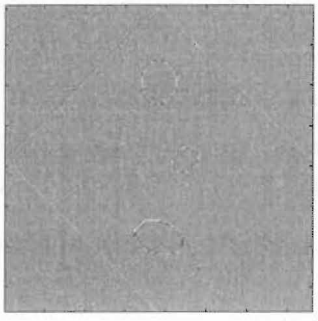
(b) Non-separable wavelet transform detail at scale = 1.



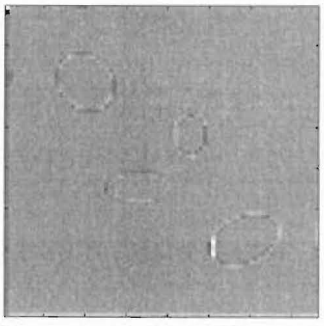
(c) Non-separable wavelet transform detail at scale = 2.



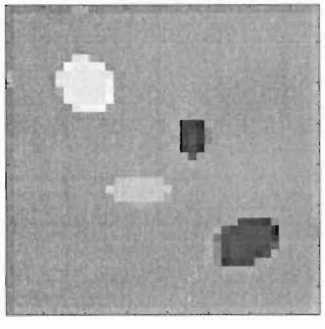
(d) Reconstruction of ellipse test figure from non-separable wavelet transform.



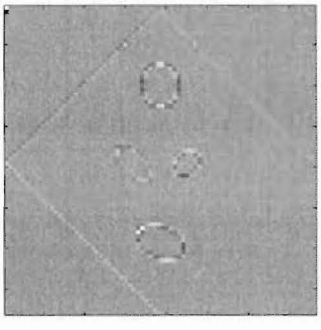
(e) Non-separable wavelet transform detail at scale = 3.



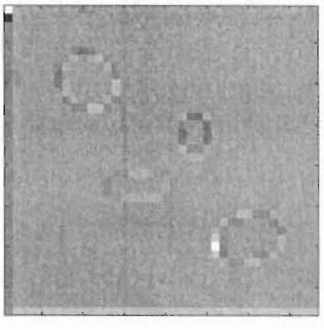
(f) Non-separable wavelet transform detail at scale = 4.



(g) Non-separable wavelet transform approximation at scale = 6.



(h) Non-separable wavelet transform detail at scale = 5.



(i) Non-separable wavelet transform detail at scale = 6.

Figure 6.6 Non-separable wavelet transform of ellipses. Left-hand column: (a) Original, (d) Reconstruction, (g) Approximation at scale, $j = 6$. Centre column: (b,e,h) Details for odd scales, $j = [1, 3, 5]$. Right-hand column: (c,f,i) Details for even scales, $j = [2, 4, 6]$.

Figure 6.8 shows the separable wavelet transform to three scales. The horizontal, vertical and diagonal biases can be seen clearly in the transform. Figure 6.9 shows the non-separable wavelet transform to six scales. As with the ellipse test figure, the orientation bias of the separable wavelet transform is removed. Certain features are isolated at different scales—for instance, the tiles midway up the picture at either side and the struts of the rose window are clear at scale 2, and the small arches at the bottom of the image are clear at scale 3.

The third type of image that has been analysed is an ultrasound image. The image in Figure 6.10 is 256x256 in size and is one of a series of images taken of a phantom of the female genital tract. A sketch of the phantom is shown in Figure 1.6. A substance known as *Echovist* has been injected into the phantom and occupies the fallopian tubes and the uterine cavity. Echovist is a substance that can be made highly echogenic for a short period of time and is used clinically for ultrasound scanning. The image was scanned from the left-hand side. The focal zone is near to the features of interest, and beam widening is apparent at the right-hand end of the image in the form of blurring. Beam widening is typical of ultrasound and is seen as elongation of features perpendicular to the direction of the scan. The elongation appears as blurring of the features and is increased as the distance from the focal zone is increased.

There are three main features of interest in the ultrasound image of the phantom. These features are indicated in the line drawing of the ultrasound image in Figure 6.11. The bright section in the middle of the left-hand side of the image represents a cross-section through the uterine cavity. The bright section towards the bottom of the image represents a cross-section through a fallopian tube. The third feature of interest represents the wall of the uterus and is quite indistinct. It is roughly circular in shape. The clearest section of the boundary of the uterus is its lower right-hand quadrant which is beside the top half of the large dark shadow at the bottom of the picture. The left-hand edge of the uterus is visible as a band of brighter pixels at the left-hand edge of the image. The other parts of the boundary are indistinct, however, the interior of the uterus is generally darker than its surrounds.

The separable wavelet transform of the ultrasound image is shown in Figure 6.12. The uterine cavity and fallopian tube are most prominent in the details at scale 3 and in the approximation. The uterine cavity is evident only in the $d^3_{1,0}$ detail and approximation, $a^3_{0,0}$. The fallopian tube is visible in all three details $d^3_{1,1}$, $d^3_{1,0}$, $d^3_{0,1}$ and in the approximation $a^3_{0,0}$.

The non-separable wavelet transform of the ultrasound image is shown in Figure 6.13. The cross-section of the uterine cavity and the fallopian tube are seen as a bright cluster of dots in the fourth and fifth levels of the transform, Figures 6.13(f) and 6.13(h). The fallopian tube is evident in the third level of transform, shown in Figure 6.13(e), but the uterine cavity is not. The uterine cavity and fallopian tube are

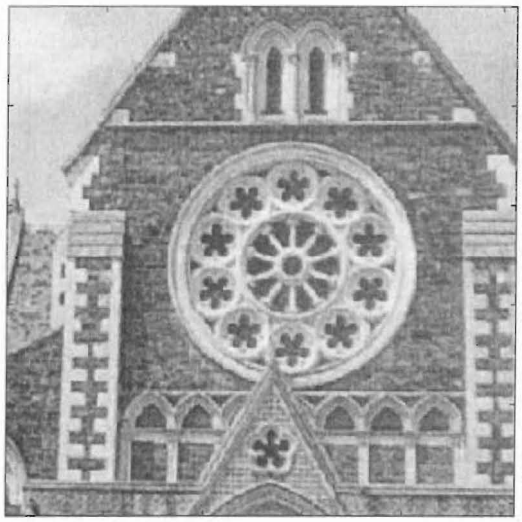


Figure 6.7 Original rose window test image.

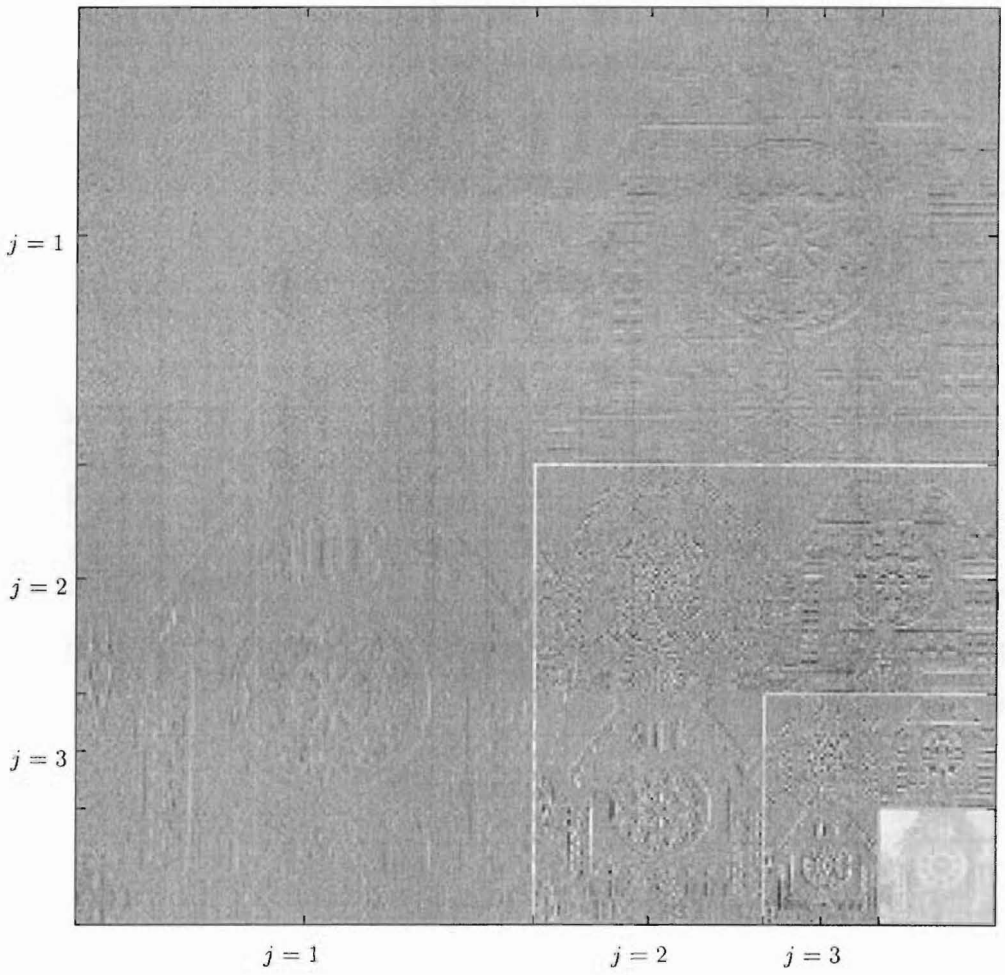
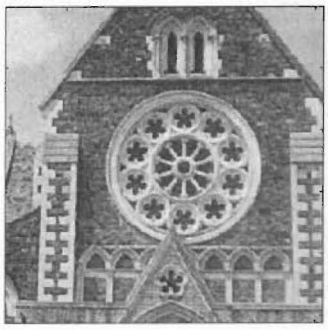
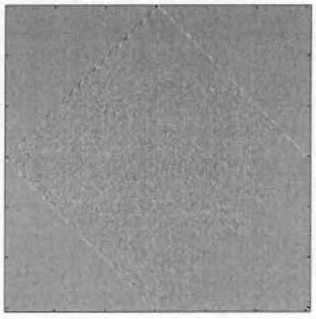


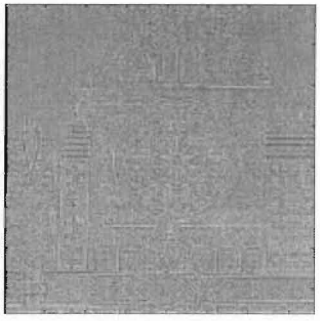
Figure 6.8 Separable wavelet transform of the rose window test image.



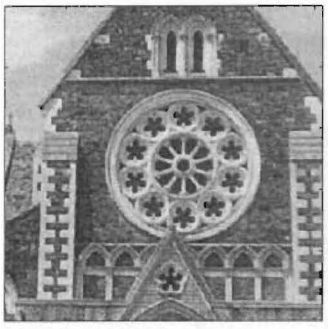
(a) Original rose window test image.



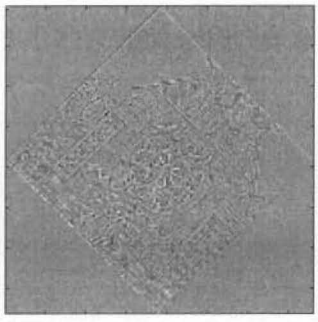
(b) Non-separable wavelet transform detail at scale = 1.



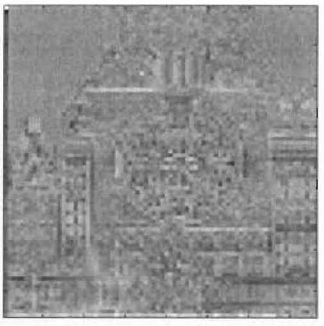
(c) Non-separable wavelet transform detail at scale = 2.



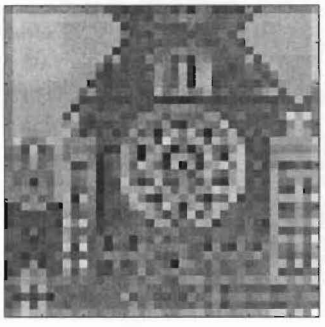
(d) Reconstruction of rose window test figure from non-separable wavelet transform.



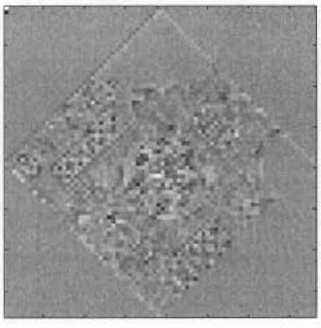
(e) Non-separable wavelet transform detail at scale = 3.



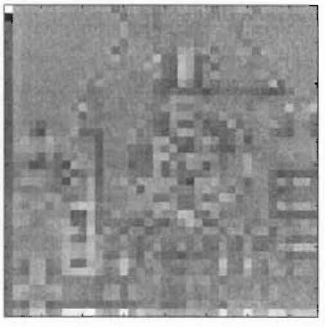
(f) Non-separable wavelet transform detail at scale = 4.



(g) Non-separable wavelet transform approximation at scale = 6.



(h) Non-separable wavelet transform detail at scale = 5.



(i) Non-separable wavelet transform detail at scale = 6.

Figure 6.9 Non-separable wavelet transform of rose window. Left-hand column: (a) Original, (d) Reconstruction, (g) Approximation at scale, $j = 6$. Centre column: (b,e,h) Details for odd scales, $j = [1, 3, 5]$. Right-hand column: (c,f,i) Details for even scales, $j = [2, 4, 6]$.

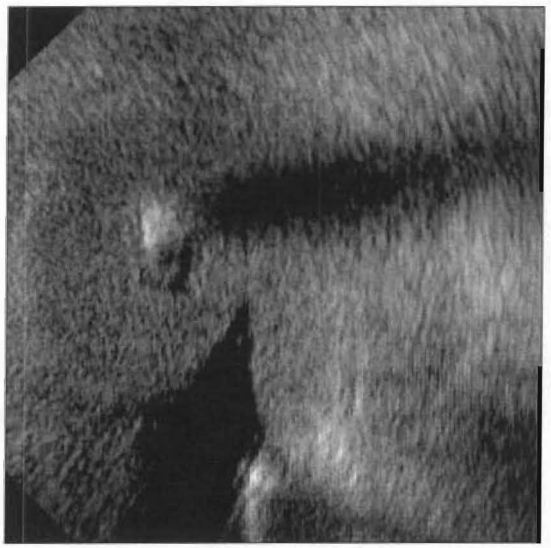


Figure 6.10 Original ultrasound image of female genital tract.

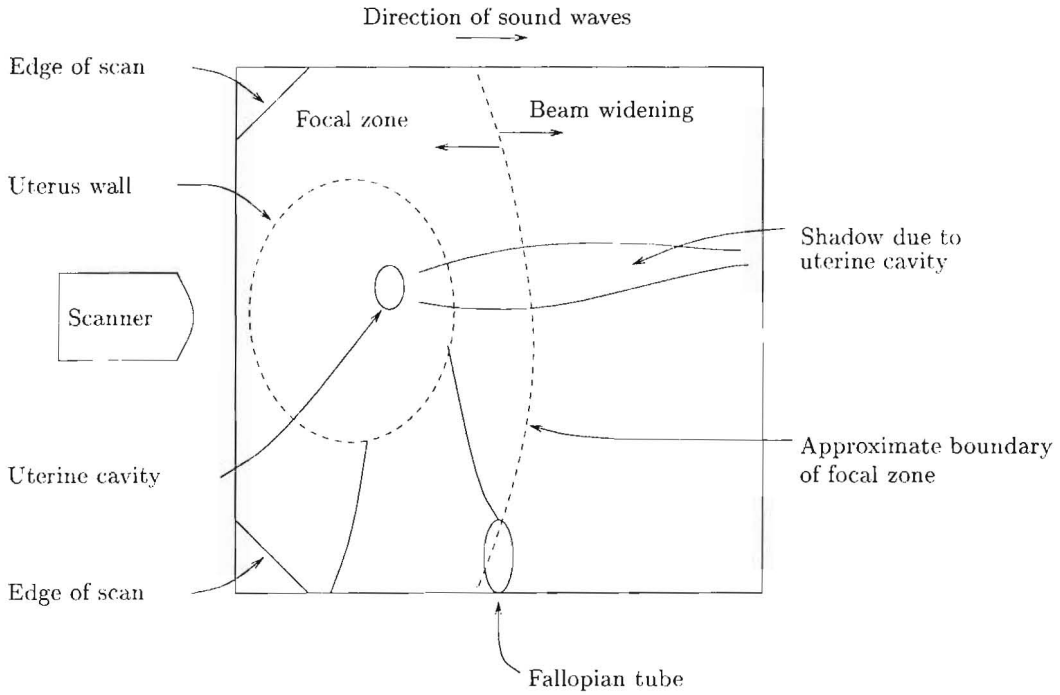


Figure 6.11 Line drawing of original ultrasound image.

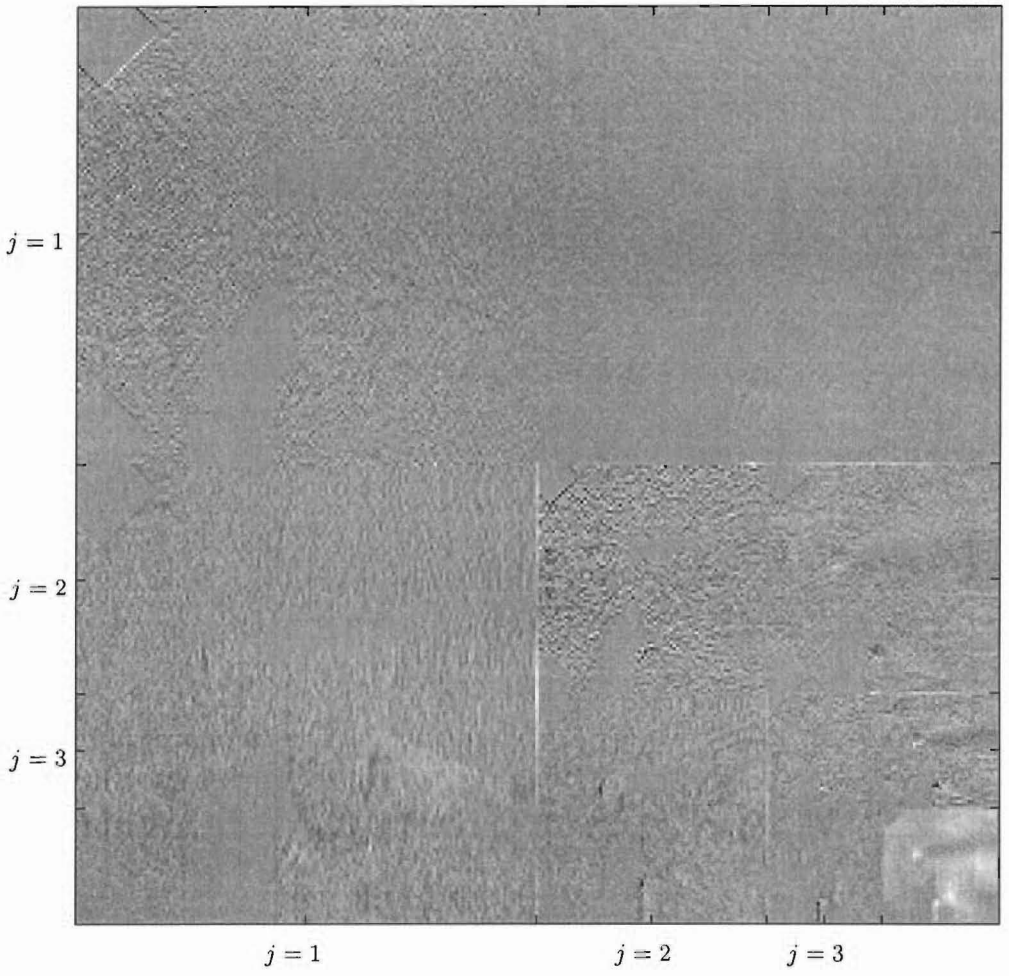


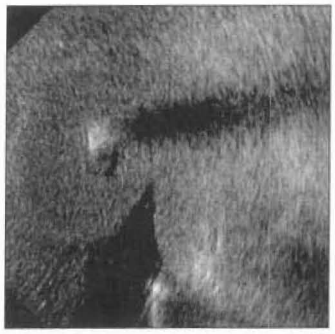
Figure 6.12 Separable wavelet transform of the ultrasound image.

not visible at scales 1, 2, or 6, but they are visible in the approximation.

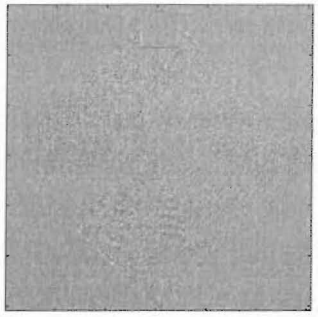
The non-separable wavelet transform using the kovalinp wavelet is shown to isolate features in the wavelet domain more successfully than the separable wavelet transform using the bior1.3 wavelet. This is due mostly to the difference in resolution between scales for the two categories. Directional components were separated in the separable wavelet transform in the ellipse and rose window images, however, for analysing ultrasound images, isolating features in different scales is more important than isolating directional components. The isolation of features in different scales is used when reducing noise and enhancing ultrasound images in Sections 8.1.3 and 8.3.

6.3.2 Characteristics of Two-Dimensional Wavelets

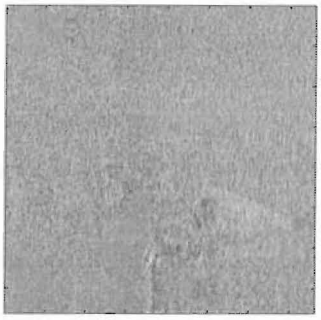
This section describes some differences between four two-dimensional wavelets and highlights some considerations for performing image enhancement. The four wavelets considered are: separable orthogonal (daub 2), separable linear phase (bior1.3), non-



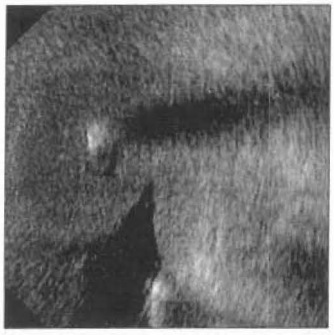
(a) Original ultrasound image.



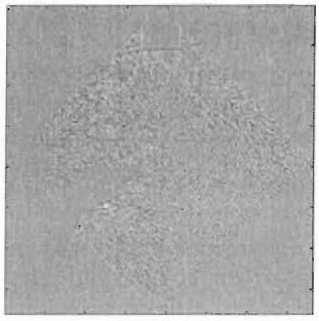
(b) Non-separable wavelet transform detail at scale = 1.



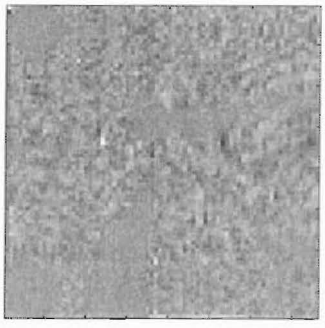
(c) Non-separable wavelet transform detail at scale = 2.



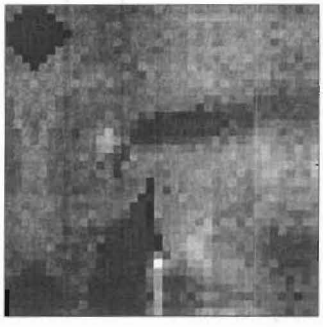
(d) Reconstruction of ultrasound from non-separable wavelet transform.



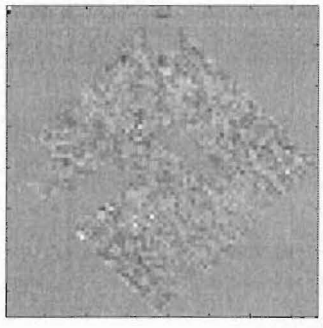
(e) Non-separable wavelet transform detail at scale = 3.



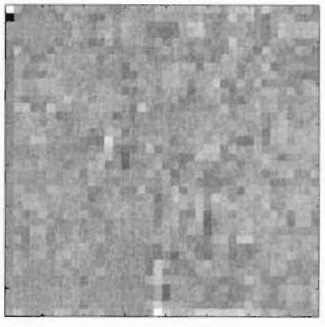
(f) Non-separable wavelet transform detail at scale = 4.



(g) Non-separable wavelet transform approximation at scale = 6.



(h) Non-separable wavelet transform detail at scale = 5.



(i) Non-separable wavelet transform detail at scale = 6.

Figure 6.13 Non-separable wavelet transform of ultrasound. Left-hand column: (a) Original, (d) Reconstruction, (g) Approximation at scale, $j = 6$. Centre column: (b,e,h) Details for odd scales, $j = [1, 3, 5]$. Right-hand column: (c,f,i) Details for even scales, $j = [2, 4, 6]$.

separable orthogonal (kovaorth, [Kovačević and Vetterli 1992], Section 6.2.2.1) and non-separable linear phase (kovalinp, [Kovačević and Vetterli 1992], Section 6.2.2.2). The two orthogonal wavelets are closely related as shown in Section 6.2.2.1. Equations (6.15), (6.16) and the two linear phase wavelets are biorthogonal of similar dimensions.

As has been described in previous sections of this chapter, the non-separable wavelets result in twice as many scales of decomposition for the same downsampling rate. This allows finer control over the denoising and enhancement processes. Ideally, the power in the image due to noise and the objects of interest is separated into different scales. The power in the decomposition at each scale, *scalar power*, $P_j(\mathbf{x})$, is calculated using a two-dimensional extension of the power equation given in Roy *et al.* [1999]:

$$P_j(\mathbf{x}) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \|X_{j,k_1,k_2}\|^2, \quad j \in \mathcal{Z}, \tag{6.25}$$

where X is the wavelet transform and $\mathbf{k} = (k_1, k_2)$ is displacement in two dimensions. To examine the characteristics of the four wavelets under consideration, the image of a noisy version of the rose window (see Figure 6.14) was decomposed by all four wavelets and the scalar power calculated. The scalar power for each wavelet is shown in Figure 6.15.

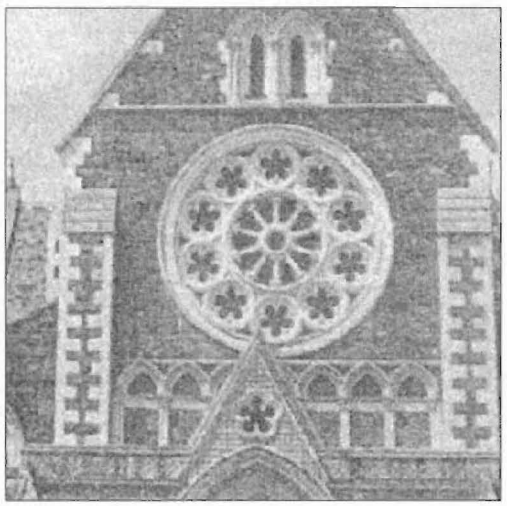
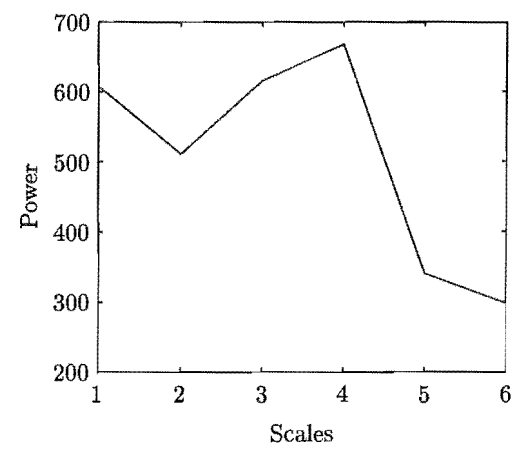
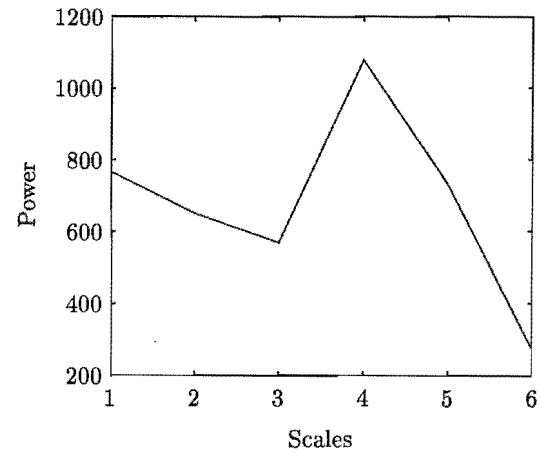


Figure 6.14 Rose window image with additive gaussian noise (mean = 0, variance = 0.005).

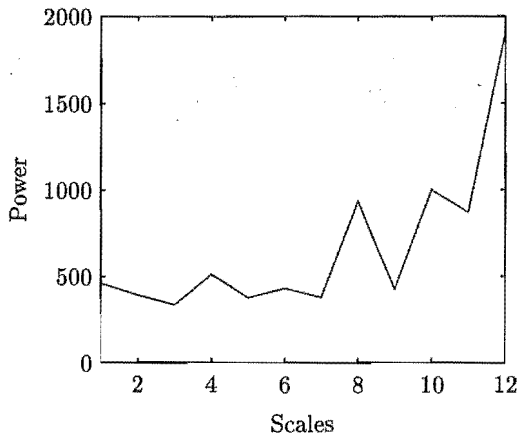
The wavelets exhibit different patterns of scalar power. The two separable wavelets, shown in Figures 6.15(a) and 6.15(b), exhibit high power in the first scale, followed by a dip, then a peak at the fourth scale. The non-separable orthogonal wavelet, Figure 6.15(c), shows a relatively constant level of power for the lower scales, rising to a peak at a scale of twelve. The non-separable linear phase wavelet, Figure 6.15(d)



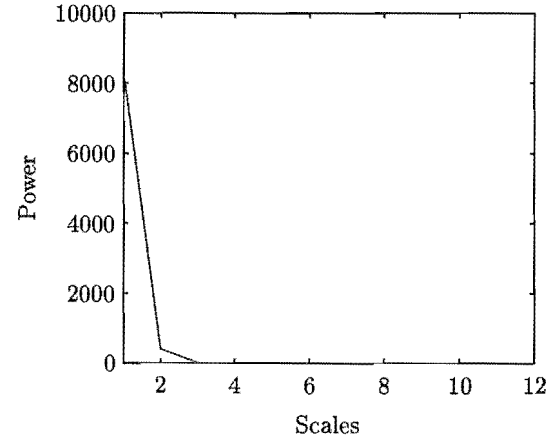
(a) Separable, orthogonal decomposition



(b) Separable, linear phase decomposition



(c) Non-separable, orthogonal decomposition



(d) Non-separable, linear phase decomposition

Figure 6.15 Power at each scale for the different wavelet decompositions (a) Separable orthogonal wavelet decomposition, (b) Separable linear phase wavelet decomposition, (c) Non-separable orthogonal wavelet decomposition, (d) Non-separable linear phase wavelet decomposition.

shows a very high concentration of power in the first scale. Of the orthogonal and the linear phase wavelets tested here, the orthogonal wavelets showed the most similarity in the scalar power.

6.3.3 Discussion

The non-separable wavelet decompositions have an advantage over the separable wavelets due to finer bands in the frequency domain which results in finer classification of features in the wavelet domain. For noise reduction and contrast enhancement, two methods of image enhancement used in Chapter 8, it is important for features to be isolated to different frequency bands, or equivalently, to different scales. This allows

noise to be (mostly) separated from image objects and each to be appropriately reduced or enhanced.

The differences in distribution of power over the scales for the different wavelets indicate that the choice of wavelet may be more important than the type of decomposition. While the non-separable wavelets result in twice as many scales of decomposition for the same downsampling rate as separable wavelets, it can be seen, when comparing the results for scalar power of the kovaorth and kovalinp wavelets, that the choice of wavelet is still important . Since almost all the power is concentrated in the first scale for the kovalinp wavelet, this decomposition would require careful thresholding when performing noise reduction and contrast enhancement of the image. In fact, performing noise reduction or contrast enhancement on this decomposition might be almost the same as on the original image. Conversely, the kovaorth wavelet has power spread fairly evenly across scales.

Other factors such as shift variance and regularity could influence a choice of wavelet [Mojsilovic *et al.* 1998]. However, as was seen in Chapter 4, measures of the wavelet characteristics may not be as useful as measures in the wavelet domain which take into consideration the interaction between the wavelets and the signal to be analysed. The next chapter investigates an image quality equation which could be used to evaluate the performance of various wavelets.

This section has illustrated differences in the scalar power of four wavelets. The non-separable wavelets result in finer frequency bands. However, care must still be taken in selecting a wavelet as individual wavelets may have undesirable frequency responses which may override the advantages of the finer frequency bands. The amount of difference in the scalar power was somewhat surprising and suggests, for instance, that appropriate methods for selecting threshold levels must be used when designing enhancement techniques.

6.4 SUMMARY

In this chapter, theory for constructing two-dimensional wavelets and some features of images in the wavelet domain with characteristics of specific wavelets were presented.

Two-dimensional wavelets can be classified as either separable or non-separable. Separable wavelets may be constructed using tensor products of one-dimensional wavelets, and non-separable wavelets may be constructed using cascade structures. Both the separable and non-separable wavelet construction techniques are simple for small filter sizes. However, solving for the non-separable case becomes non-trivial as the filter sizes increase. Separable and non-separable wavelets that are used in Chapter 8 were presented.

Features of images were isolated to different scales more successfully using a non-

separable wavelet compared with using a separable wavelet. Main directional components, horizontal, vertical and diagonal, were isolated using a separable wavelet. However, since the main interest here is ultrasound images, where the need for directional analysis is small and the need for feature extraction is large, it is likely that a non-separable wavelet will be more appropriate for analysis.

The scalar power of four wavelets were evaluated and compared. The four wavelets were: an orthogonal separable wavelet, constructed using daub2; a linear phase separable wavelet, constructed using bior1.3; an orthogonal non-separable wavelet, kovaorth; and a linear phase non-separable wavelet, kovalinp. The orthogonal wavelets were closely related and the linear phase wavelets were chosen to be a similar size. Using scalar power as a criteria, the choice of wavelet seems more important than the type of decomposition. However, for the two-dimensional daub2 wavelet and the kovaorth wavelet, the category of wavelets that have the most similar distributions in scalar power, the non-separable wavelet provided the most flexibility for enhancement schemes. The performance of these wavelets is compared in Chapter 8.

Chapter 7

IMAGE QUALITY

Enhancement of images necessarily raises the question of how to assess image quality. Assessing image quality is a subjective issue. The task is complicated by the many factors, such as contrast in the image, sharpness of objects, noise and artefacts. People have their own preferences that are not necessarily consistent, but may vary depending on the type of image being assessed, or the subject of the image. Therefore, to assess the quality of the images obtained by image enhancement, a more objective process is desired. A process based on measurable characteristics of an image has the potential to automate the assessment process, and ultimately, to lead to a system that can iteratively improve the quality of an image.

There are many different approaches in the literature to objectively assessing image quality, and the next section describes some of these approaches. However, none of these approaches directly address assessment of image quality in ultrasound images. Therefore, in Section 7.2.1, an Image Quality Equation is developed for assessing ultrasound images. The equation is tested in Section 7.2.2, and discussed in Section 7.2.3.

7.1 EXISTING IMAGE QUALITY MEASURES

Image quality measures are used in a variety of settings, from commercial to medical applications, aerial photography to satellite imagery. This section reviews a number of methods for measuring image quality in the literature. The relevance of these methods to measuring the quality of ultrasound images is discussed at the end of the section.

Image quality for digital libraries and archives is defined using standards based on micrographics [Kenney and Chapman 1996], and the Quality Index (QI). The relationship between text legibility and system resolution is described with the QI. The QI is found by multiplying the height of the smallest significant character h by the smallest line pair pattern resolved by a camera on a technical test target, p.

Digital resolution is commonly measured in dots per inch (DPI) and classical resolution is measured in lines pairs per millimeter. To determine classical resolution from digital resolution, multiply DPI by 0.039 to get dots per millimeter. Line pairs per

millimeter (one black and one white line) are represented by two dots, so DPI must be divided by 2 to get an equivalent measurement. In order to compensate for possible image detector mis-alignment, the input scanning resolution is increased by at least 50%. Therefore, classical resolution in line pairs per millimeter is given by (0.039 x DPI) / 3, and QI is given by 0.013 DPI x h. For a DPI of 300, and a text that has characters at least 2mm high, the QI is 8, or excellent rendering. The reference for the QI has been established as 8 = excellent, 5 = good, 3.6 = legible, 3 = poor.

In a study of the quality of colour images [Engeldrum 1995], image quality is defined using a generalised weighted mean hypothesis. It is proposed that weighted attribute measures can be combined in the following manner:

$$Image\ Quality = a_0 + (a_1 attribute_1^p + attribute_2^p + ...)^{\frac{1}{p}}, \tag{7.1}$$

where a_k and p are constants and $attribute_k$ are measures of image quality components such as darkness, sharpness, and graininess. The attributes are defined as *customer perceptions*, which were subjectively measured. The constants were calculated using linear regression on results from a group of people ranking a test set of images for several attributes. The attributes found to have statistically significant regression coefficients were *sharpness, defects* and *colour accuracy*, which are defined by:

- Sharpness refers to the distinction of boundaries between colours;
- Colour accuracy refers to how well a colour print reproduces the colour in an original 35-mm slide;
- A defect is any imperfection in a print that is not in the original slide.

Two other attributes found not to have statistically significant regression coefficients were *colour uniformity* and *colourfulness*, defined by:

- Colour uniformity refers to the consistency of the colour in uniform areas of colour;
- Colourfulness describes how bright and saturated the color in the print looks.

Engeldrum [1995] concludes that image quality, as perceived by the consumer, is a generalised average of the customer perceptions, or equivalently, that image quality may be viewed as a generalised distance from some image quality origin.

Response times of humans to recognise objects has been used as a measure of quality in progressive image transmissions (i.e. coarse scales followed by progressively finer scales) for JPEG, a progressive version of JPEG, and two wavelet zerotree methods:

Embedded Zerotree Wavelet (EZW) coding due to Shapiro and Set Partitioning in Hierarchical Trees (SPIHT) due to Said and Pearlman [Cen *et al.* 1998].

The quality of compressed magnetic resonance and computed tomography images has been evaluated using signal to noise ratio (SNR), subjective rating and diagnostic accuracy [Cosman *et al.* 1993, Cosman *et al.* 1994, Gray *et al.* 1995, Gray *et al.* 1996, Perlmutter *et al.* 1997]. The original images are taken as the "gold standard" with the images rated by a number of radiologists for contrast, sharpness, motion unsharpness, and the number of correct diagnoses, including counting missed diagnoses and false diagnoses. The SNR is given by:

$$SNR = 10log_{10}\frac{\sigma_g^2}{E\left[\|g-\hat{g}\|^2\right]}, \tag{7.2}$$

where $E\left[\|g-\hat{g}\|^2\right]$ is the average distortion when comparing the original image g with the compressed image, $\hat{g}$. The subjective rating was based on experts rating each image according to "its usefulness for the measurement task." Diagnostic accuracy was measured against the diagnosis of the original. Both sensitivity, the probability that something is detected given that it is present in the original, and predictive value positive, PVP, the chance that an abnormality is actually present given that it is marked, were used to determine diagnostic accuracy.

Veldhuizen [1998] presents mean squared error (MSE) and peak signal to noise ratio (PSNR) as standard methods for measuring image quality. The MSE, e_{MSE}, between two images $g(x,y)$ and $\hat{g}(x,y)$ of size MxN is:

$$e_{MSE} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[\hat{g}(n,m)-g(n,m)\right]^2. \tag{7.3}$$

The MSE depends heavily on the number of bits used to represent each pixel. To avoid this dependence, the PSNR is used, which scales the MSE according to the intensity range:

$$PSNR = -10log_{10}\frac{e_{MSE}}{S^2}, \tag{7.4}$$

where S is the maximum pixel value. The main failing of this measure is that the signal strength is estimated as S^2, rather than the actual signal strength. The PSNR is good for comparing intra-image results, but not inter-image results [Veldhuizen 1998]. The PSNR offers no information regarding which frequencies of the signal may have been lost.

To isolate different frequency components that might be affected, frequency domain SNR behaviour is also considered. The discrete Fourier transform of an image $g(n,m)$

is given by:

$$G(u,v) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} g(n,m)\, e^{-2\pi j\left(\frac{un}{N}+\frac{vm}{M}\right)}. \tag{7.5}$$

It is useful to group coefficients according to normalized spatial frequency, λ^{-1}:

$$\lambda^{-1} = \sqrt{\frac{N^2v^2 + M^2u^2}{M^2N^2}}, \tag{7.6}$$

which has a range of $[0, \frac{1}{\sqrt{2}}]$. λ is the wavelength of the spatial frequency component. Therefore, features on the scale of 10 pixels ($\lambda \approx 10$) have a spatial frequency of $\lambda^{-1} \approx 0.1$. The DC component or average of an image corresponds to $\lambda^{-1} = 0$. The maximum frequency is $\lambda^{-1} = \frac{1}{\sqrt{2}}$.

For an original image, $g(n,m)$, and degraded version, $\hat{g}(n,m)$, the difference image is:

$$d(n,m) = g(n,m) - \hat{g}(n,m). \tag{7.7}$$

The Fourier transforms of the original image, $G(u,v)$, and difference image, $D(u,v)$, are taken and the frequency domain separated into non-overlapping frequency bands, $B_i = [\lambda_i^{-1}, \lambda_{i+1}^{-1})$. The signal to noise ratio for each band is given by:

$$SNR(B_i) = \frac{P_G(B_i)}{P_D(B_i)}, \tag{7.8}$$

where the average band power, $P_G(B_i)$, is calculated by averaging $|S(u,v)|^2$ over all (u,v) and the average band power for the difference image, $P_D(B_i)$, is calculated similarly.

It is has been shown that simple approaches such as mean squared error do not provide meaningful measures of image fidelity [Taylor *et al.* 1997]. Some research suggests that a model of the human visual system (HSV) based on discrimination thresholds (rather than detection thresholds) appears to be more psychophysically plausible and that this model can by used to assess image fidelity [Taylor *et al.* 1997].

Comes *et al.* [1995] use masks based on the human visual system to distinguish areas that may be considered background from those that may be considered foreground. The noise is estimated in the foreground areas. This estimation of noise is used in the determination of a masked peak signal to noise ratio ($MPSNR = 255^2/P_B$) and a *quality factor*, $Q = 5/(1 + KP_B)$, where P_B is the mean squared error of the noise in the foreground areas and K is a normalisation factor to give Q between 1 and 5. The initial masks are based on *perceptual channels* and these channels, based on small portions of the image, vary depending on the orientation and spatial frequency of the

portions.

Westen *et al.* [1995] developed a perceptual error measure also based on the human visual system. Their measure is based on local band-limited contrast for different frequencies and orientations.

Heeger and Teo [1995] use retinal, cortical and detection models to determine a probability map of the image. The map measures the probability that a viewer would experience visual distortion at each position in the image. An average value is given as the overall image fidelity measure.

Researchers working with aerial images have devised an image quality measure (IQM) based on the power spectra of the image [Nill and Bouzas 1992]. This technique assumes that power spectra of images is invariant from scene to scene and from scale to scale. For an MxM two-d image, with a pixel located at (x, y), and intensity, $h(x, y)$, the discrete FT of the image $H(u, v)$ is:

$$H(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} exp\left[-2\pi i y \frac{v}{M}\right] exp\left[-2\pi i x \frac{u}{M}\right] h(x,y), \tag{7.9}$$

where u, $v = -M/2 \ldots M/2$ are normalised to give u/M and v/M cycles per pixel width. The two-dimensional image power spectrum of $h(x, y)$ is defined as $|H(u, v)|^2$. To compensate for the effect of image to image brightness, the power spectrum is normalised using either the square of the average gray level of the image, $\mu^2 = |H(0, 0)|^2/M^4$, or the total power. μ^2 is also known as the dc power. The power spectrum is also normalised for the size of the image, M^2. The normalised power spectrum, $P(u, v)$, is given by:

$$P(u,v) = \frac{|H(u,v)|^2}{\mu^2 M^2}. \tag{7.10}$$

The IQM [Nill and Bouzas 1992] incorporates the Human Visual System (HVS) model, $A(T\rho)$, a noise filter, $W(\rho)$, and a directional scale factor, $S(\theta_1)$, and is expressed in polar co-ordinates, (ρ, θ):

$$IQM = \frac{1}{M^2} \sum_{\theta=-180^\circ}^{180^\circ} \sum_{\rho=0.01}^{0.5} S(\theta_1) W(\rho) A^2(T\rho) P(\rho, \theta). \tag{7.11}$$

The directional scale factor takes into account the pixel width, q, the height from the ground, h, and focal length of the imaging device, f, resulting in a weight measured in spatial frequency units of cycles per ground meter: $S(\theta_1) = f/2hq$. The noise filter is a modified Weiner filter and is applied in order to reduce the impact of noise on the power spectrum. The HVS model is a rotational symmetric modulation transfer function that

filters the power spectrum response given by: $A^2(T\rho) = (0.2 + 0.45T\rho)exp(-0.18T\rho)$, where T is a constant and fixes the spatial frequency of the peak of the function with respect to the image's Nyquist frequency, and $T\rho$ is in units of cycles/degree subtended by the eye.

Researchers working with military images describe a general image quality equation (GIQE) [Leachtenauer *et al.* 1997, Driggers *et al.* 1998]. This equation has been developed to predict the level to which objects within an image can be interpreted. The standard to which the levels are tested against is the military National Imagery Interpretability Rating Scale (NIIRS). The NIIRS rates an image in terms of being able to distinguish different size objects and classify them. Rating 0 is for a very poor image. Rating 1 implies a medium sized port facility may be detected or that taxiways and runways may be distinguished at a large airfield. The top end of the scale, rating 9, implies vehicle registration numbers may be identified on trucks, or that cross-slot heads on aircraft skin panel fasteners are differentiable from single slot heads.

The GIQE predicts the NIIRS using scale, resolution, sharpness, contrast and noise. Scale and resolution are accounted for by using the ground sampled distance, GSD, given by:

$$GSD = \frac{\left[\left(\frac{\text{pixel pitch}}{\text{focal length}}\right) * \text{slant range}\right]}{\cos(\text{look angle})}. \tag{7.12}$$

Sharpness is measured using the relative edge response, RER, which is the slope of the system's edge response and is derived from the modulation transfer function, MTF. In practice, RER is measured between two points that are 0.5 pixels from an edge, normalised and the geometric mean is taken to give RER_{GM}. The gain, G, due to modulation transfer function compensation, MTFC, or edge sharpening, given by the root of the sum of the squares of the MTFC kernel values, is accounted for as well as the geometric mean height overshoot caused by edge sharpening, H_{GM}. The SNR is also a component of the GIQE and is a measure of the contrast [Leachtenauer *et al.* 1997]. The GIQE is given by:

$$\begin{aligned} NIIRS = 10.251 - a\log_{10} GSD_{GM} + b\log_{10} RER_{GM} - (0.656 * H_{GM}) \\ - (0.344 * G/SNR), \end{aligned} \tag{7.13}$$

where $a = 3.32$, $b = 1.559$ for $RER \geq 0.9$, and $a = 3.16$, $b = 2.817$ for $RER < 0.9$. The constants have been determined by regression on a large number of test examples ranked by human experts.

None of the above approaches can be applied directly to ultrasound images. The Quality Index for commercial applications applies to the quality of text in an image. Text typically has a high contrast relative to the background and the imaging processes

introduce relatively small errors. Ultrasound images do not have a high contrast between the objects and background and also have more artefacts and noise introduced by the imaging process. However, the concept of object size in the image is a useful one and is also used in the aerial image equations. The aerial image equations use parameters, such as the height from the ground and angle the image was taken at, to estimate the size of objects that will be discernable. These parameters have no relevance for ultrasound imaging. Several image quality measures use parameters based on the difference between the original image and a second one that has been transformed from the first. This approach is not appropriate here as the original image contains noise and artefacts and in the transformation of this image, there is no guarantee that the image will be improved, therefore the difference between the new image and the original may be measuring either the addition or removal of noise or artefacts. It is beyond the scope of this thesis to investigate the human visual system. The power of the spectra of the images and the RER may be able to be applied. Since none of the techniques presented in this section are able to be applied directly to ultrasound images, a new image quality equation, based on the techniques presented in this section, is developed.

7.2 NOVEL IMAGE QUALITY EQUATION

This section describes the development and testing of a novel image quality equation. The new equation is developed to be applied to ultrasound images. A discussion of the performance of this image quality equation is also presented.

7.2.1 Development of Image Quality Equation

The images to be assessed for image quality in this thesis are ultrasound images. Compared to most other medical images, they are inherently of poor quality, containing high levels of noise, artefacts, and low resolution. If these images are subjectively judged to assess their image quality, any person grading an image must trade off noise, contrast, and sharpness among other factors. Whilst one image might have sharply defined objects, it might also have a large amount of noise. If this image is to be compared to an image of a similar subject with little or no noise, but with a large amount of blurring, then personal preference makes consistent assessment difficult, even for the same observer. Therefore, it is important to use objective measures so that consistency can be obtained. This section develops a novel image quality equation to be used specifically for ultrasound images. The features to be measured are discussed.

There are several features of an image that were identified as being important in image quality. These features were identified through discussion with a radiologist [Anderson 1994], as well as reviewing the literature. Important features are: object size

within the image, edge sharpness, contrast, magnitude of noise, and frequencies of noise. Each of these features is considered in turn.

The QI [Kenney and Chapman 1996] effectively uses the size of the smallest object and the resolution to give an image quality assessment. The IQM [Nill and Bouzas 1992] and GIQE [Leachtenauer *et al.* 1997, Driggers *et al.* 1998] use a range component in their image quality equations. The QI deals with text and the IQM and GIQE both assess aerial images; an equivalent measure is required for ultrasound images assessed in this thesis. The size of an object in an image may be defined by the number of pixels the object covers. Therefore, if the resolution is expressed in mm per pixel, an equivalent metric to the QI, the Pixel Quality Index, PQI, can be expressed by:

$$\text{PQI} = \frac{w\ 0.039\ DPI\ I_{mm}}{I_p} \tag{7.14}$$

where w is the width of the object in pixels, DPI is the image resolution in dots per inch, I_{mm} is the image size in millimeters, and I_p is the image size in pixels. The QI resolution is measured in line pairs per millimeter and accounts for a mis-alignment by an image detector. Since neither of these are of importance here and are constant, we can adjust the reference values for the QI. Therefore, the following reference can be used: 24 = excellent, 15 = good, 10.8 = legible/perceivable, 9 = poor. So, for example, an image 256x256 pixels, printed at 600 DPI and 60x60 mm would have a basic image resolution of 5.48 dots/pixel. For an object to be very clear in this image (excellent quality, QI=24), it would need to cover 24/5.48 = 4.4 or 5 pixels. The object would have good clarity if it was 3 pixels wide, (PQI = 16.44). If an object covered 2 pixels, PQI = 10.96, it would be perceivable. However, an object 1 pixel wide would have poor visual clarity.

One problem with the PQI is that for noisy or blurred images, objects often appear larger than in the original images. Therefore, for an original image with the same number of pixels, size and DPI as the example above, if the smallest object to be recognised is 2 pixels wide, then the PQI is 10.96. However, for a blurred version of the same image, if the smallest object that can be recognised is 5 pixels wide, then the PQI would be 27.4. This is clearly not ideal as the blurred image has a better PQI than the original image. Since the PQI should have a linear relationship with the image quality, to rectify the problem, the PQI is multiplied by a ratio to give the adjusted pixel quality index, PQI_{adj}:

$$PQI_{adj} = \frac{w_t}{w_a} PQI. \tag{7.15}$$

where w_t is the true width, in pixels, of an object as it would appear in an image with no noise or artefacts, and w_a is the actual width, in pixels, of the object in the image.

For the ultrasound images used later, the true width of the object is estimated using a priori knowledge of the physiology.

Edge sharpness is measured in pixels using the normalised geometric mean proposed for use in the GIQE [Leachtenauer *et al.* 1997, Driggers *et al.* 1998], the *RER*, which is the slope of the system's edge response (see Section 7.1). In practice, *RER* is measured between two points that are 0.5 pixels from an edge, normalised and the geometric mean is taken to give RER_{GM}. The *RER* is also a measure of contrast.

The one-dimensional conversion of the two-dimensional power spectrum can be used to assess the frequencies at which the noise occurs. (See Equations (7.10) and (7.9).) As seen in Nill and Bouzas [1992], the power spectrum of an image with noise has a much flatter response at the high frequency end compared with the power spectrum of the same image with no noise. A denoised image has lower relative power at the high frequency end of the spectrum. This is illustrated in Figure 7.1. The sum of the power at all frequencies will, therefore, be lower for an image without noise compared to the same image with noise. This difference is used as an indication of noise reduction.

Two different approaches to combining these values have been used previously by researchers. Linear regression has been used [Engeldrum 1995, Leachtenauer *et al.* 1997, Driggers *et al.* 1998]. Alternatively, a sum of the products of different effects over a range of parameters [Nill and Bouzas 1992] has been employed.

The best results seem to have been obtained using regression. Therefore, this approach is used, leading to an equation of the form:

$$IQ = a_0 + a_1 PQI_{adj} + a_2 RER + a_3 P, \tag{7.16}$$

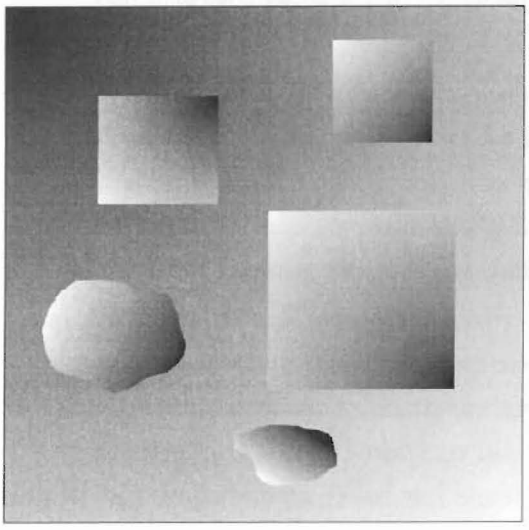
where P is the sum of the power of the image. The equation is tested in the next section.

7.2.2 Testing the Image Quality Equation

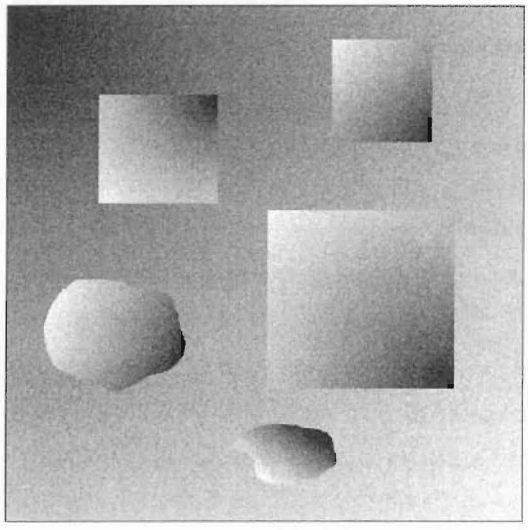
The image quality equation was tested using a subset of the set of the phantom ultrasound images introduced in Section 1.3. The aim of testing is to ensure that the equation does indeed measure image quality.

One of these images was manipulated to produce images of varying quality which were ranked and used as a reference. A set of six different images was similarly manipulated, the image quality equation parameters evaluated, and the images graded by six people. The grading was performed twice to test for intra-operator variance. The results of the gradings were used in a regression, resulting in an image quality equation. Another set of six images was manipulated to verify the image quality equation. This process is described in more detail below.

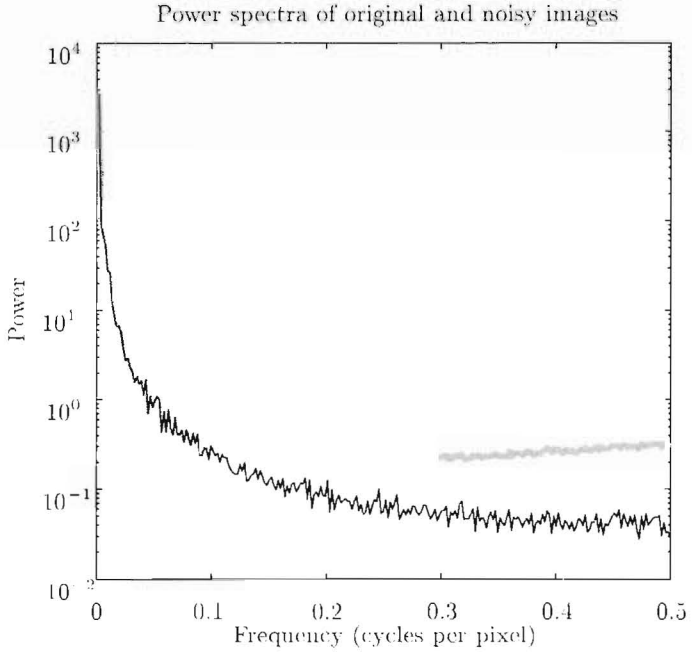
A set of 54 ultrasound images were available for use in image enhancement and testing. The ultrasound images were acquired using an Acuson scanner of a phantom of the female genital tract. The 54 images are slices that were taken in one sweep of



(a) Original test image



(b) Test image with noise



(c) Power spectrum of original test image (black line) and noisy test image (grey line).

Figure 7.1 (a) Test image and (b) noisy counterpart with (c) power spectrum of both (a) and (b), illustrating the different amounts of power at different frequencies for both images.

the phantom. Thirteen images distributed throughout the set of 54 ultrasound images were used for testing the image quality equation. These are referred to as the original images.

One original image was manipulated to change the contrast and blurring in the image. The manipulation was performed for five different settings. Some of these

images were judged worse, and some better than the original image. The original image was not used as part of the grading process. The manipulated images were then ordered and ranked 1 to 5 (1=worst, 5=best.) These images are referred to as the graded images and were supplied as a reference guide for people grading another set of images.

A set of six different original images were manipulated to change the contrast or blur the images in the same manner as the graded images. This resulted in 30 different images (not including the originals). These are referred to as the ranked images.

The 30 ranked images were randomly ordered and divided into two groups of 15. This process was repeated three separate times, resulting in six sets of 15 images, randomly ordered. The images were randomly ordered to reduce learning effects [Cosman *et al.* 1994]. Each group of images was given to a different person to assess. The individuals had a variety of experience with medical images. Each person was instructed to grade the images based on the following assessment criteria:

- ability to accurately define positions of features of interest, such as fallopian tubes, uterine cavity or uterus wall,
- contrast in the image, particularly of objects of interest against the background,
- lack of noise or other artefacts,
- sharpness of the image.

Each person was instructed to rank the images with a number between 0 and 5, according to these criteria, and using the graded images as a guide. At an interval of no less than three days, each person repeated the task with the same group of images. One person graded a different group of images, but seven of the 15 images were the same.

The image quality equation parameters, RER, adjusted PQI and P, were evaluated for each of the ranked images. These parameters were used as the independent variables in a regression analysis, with the grades given by the six people from the first test used as the dependent variable. The regression was performed in Excel 2000, resulting in the following equation:

$$IQ = 3.7 - 1.7RER - 0.065PQI + 49P \tag{7.17}$$

The statistics for the regression are shown in Table 7.1 and indicate that the equation is a poor fit to the data ($R^2 = 0.29$). This could be due to a number of factors. The first is that each person ranking the images may not have been consistent. The second is that there may have been a large variation between people in their

rankings. A third possibility is that the IQE parameters do not measure image quality as expected.

Multiple R	0.54
R Square	0.29
F	3.6
Significance F	0.028
p values:	
Intercept	8.3E-06
RER	0.52
PQI(adjusted)	0.27
P/N^2	0.16

Table 7.1 Statistics for the image quality equation regression.

The first possibility was tested by comparing the grades from each person from the two ranking tasks. The results should be close to identical if the observers are consistent. For every image each person graded, the difference between the two grades was taken. The mean and variance were calculated for the difference and the magnitude of the difference. The results are shown in Tables 7.2 and 7.3, respectively. D and E were remarkably consistent with only 3 and 2 changes, respectively. A was also fairly consistent. Even though A changed many rankings (13), the changes were small (mean difference = 0.14, variance = 0.24). However, B, C, and F did not have very consistent rankings (12, 11 and 4/7 changes, respectively, all with mean difference > 0.5) . There was a variation of the rankings in the retests, indicating intra-person differences exist. In the retest, there was no pattern to the changes in the rankings. Most people had changes in both positive and negative directions. There were no clear patterns to the changes for individual images.

Person	A	B	C	D	E	F
Number of different rankings	13	12	11	3	2	4/7
Mean difference	0.14	-1.1	-0.55	1	1	1.3
Variance of difference	0.24	0.60	0.62	0	0	2.9

Table 7.2 Number of differences, mean of the differences, and variance of differences in rankings (Ranking 1 - Ranking 2).

Person	A	B	C	D	E	F
Number of different rankings	13	12	11	3	2	4/7
Mean difference	0.44	1.2	0.82	1	1	1.8
Variance of difference	0.054	0.38	0.21	0	0	0.92

Table 7.3 Number of differences, mean of the magnitude of differences, and variance of magnitude of differences in rankings (|Ranking 1 - Ranking 2|).

To test whether there was a large variation between people, the regression was performed for each person for the ranks from the first ranking. The regression statistics

are shown in Table 7.4. There is a large variation between people in their regression results, indicating significant inter-personal differences.

Regression Statistics	A	B	C	D	E	F
Multiple R	0.60	0.45	0.80	0.40	0.31	0.66
R Square	0.36	0.20	0.63	0.16	0.10	0.44
Adjusted R Square	0.19	-0.013	0.53	-0.068	-0.15	0.29
Standard Error	1.0	1.1	0.82	1.4	0.98	1.2
Observations	15	15	15	15	15	15
F	2.1	0.94	6.4	0.71	0.40	2.9
Significance F	0.16	0.46	0.009	0.57	0.76	0.085
p-value intercept	0.013	0.027	0.000	0.015	0.077	0.001
p-value RER	0.19	0.40	0.42	0.99	0.78	0.63
p-value PQI	0.97	0.70	0.017	0.60	0.64	0.45
p-value P/N^2	0.60	0.26	0.023	0.48	0.81	0.11

Table 7.4 Regression statistics for individuals.

Regression of each individual's rankings revealed that most did not provide a good fit. The person with the best fit (C) had had more exposure to the type of image and enhancement procedure, suggesting that there might be an element of training missing with the other observers. The second round of rankings did not change the regression results significantly, with each person's R^2 value within 0.12 of their previous result. A, D, and F all saw a drop in R^2, while the other three all saw a rise. C's went up from 0.63 to 0.71. The regression using the means of the individual ranks for each image had an improved R^2 value of 0.30, compared to 0.29, which is still not a good fit.

Both rounds of rankings of the images did not produce a regression with a good fit. Some intra-personal and inter-personal differences were found. The significance of each of the equation parameters varied widely between individuals. This indicates that there are other unknown mechanisms at play. A fourth possibility for the poor fit is that the parameters for the IQE do not perform as expected in terms of measuring salient image quality information. Therefore, the IQE parameters were examined.

One of the independent variables in the IQE is power. It had been thought that the power would decrease for the enhanced images, but this is not the case. This discrepancy is demonstrated by considering a set of images and an approximation of the their power spectra. Plotting the spectra of the enhanced images relative to the spectra of the original image shows the difference in power relative to the frequency. Figure 7.2 shows one set of images that were ranked, and Figure 7.3 shows the corresponding spectra. Each plot of the spectra of the enhanced images has the spectra of the original image as a solid black line and the enhanced image as a grey line. The original image is shown in Figure 7.2(a). The image shown in Figure 7.2(b) has very high contrast between the uterine cavity and the rest of the image. Unfortunately, this is at the expense of intensity in the fallopian tube. This image is considered worse than the

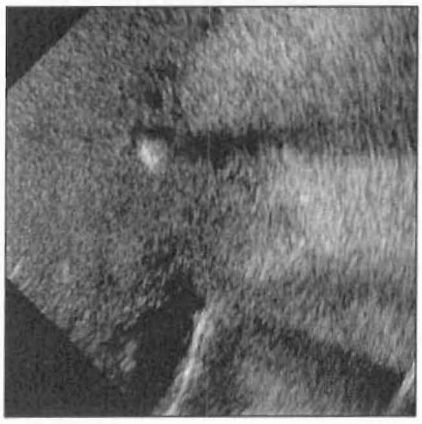
original due to the excessive contrast lost in the background. The image in Figure 7.2(c) is grainy and there is poor contrast between the uterine cavity and the fallopian tube with the background. However, the effects due to beam widening have been reduced. Figure 7.2(d) shows an image that has retained the contrast of the uterine cavity and the fallopian tubes with the background and has reduced the effects of the beam widening. Based on the criteria for assessing ultrasound images outlined previously, this image is considered better than the original. Figures 7.2(e) and (f) are blurred images of the original with Figure 7.2(f) more blurred than Figure 7.2(e).

Turning to the corresponding power spectra, the trend in the original image is for a generally decreasing spectrum from low to high frequencies, with a small increase at the highest frequencies. There is a flattening of the spectrum at about 0.05 cycles per pixel (cpp). The dark image has a spectrum (Figure 7.3(b)) that follows the spectrum of the original until 0.15 cpp, and then diverges from the original, getting larger until about 0.3 cpp where the enhanced image spectrum seems to stop diverging and maintain the distance from the original spectrum. The spectrum in Figure 7.3(c) (the grainy, low contrast image) diverges from the original at about 0.03 cpp, rising away from the original until about 0.1 cpp, then maintaining the distance away from the original. The rise from the original is abrupt, resulting in a trough in the spectrum at about 0.03 cpp and a large difference between the enhanced and original image spectras. Figure 7.3(d) displays the spectrum from the image that is better than the original. The spectrum diverges from the original even earlier than the spectrum in (b), but the divergence is not as large. The increasing divergence only lasts from about 0.02–0.05 cpp, and thereafter, the difference is maintained. In Figures 7.3(e) and (f), the spectra decrease compared with the original. The spectrum from the enhanced image in (f) decreases the most, seeming to decrease at a steady rate compared with the original between 0.1 cpp and 0.4 cpp.

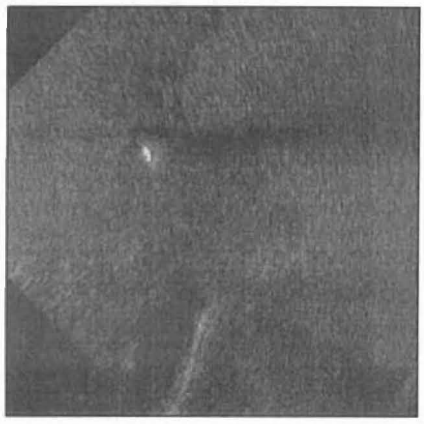
From these observations, the *shape* of the spectrum appears to be more important than the area under the curve as assumed in the image quality equation. Therefore, either a more specific measure of the power is required, or a different measure that contains some information about the noise level in the image is required.

The regression equation was tested using the another set of images. Another six original images were manipulated, each with five different settings of the enhancement algorithm, described in the following section. Some of these settings were chosen to enhance the image, while others were chosen to provide worse images. These 30 images are referred to as the verifying images.

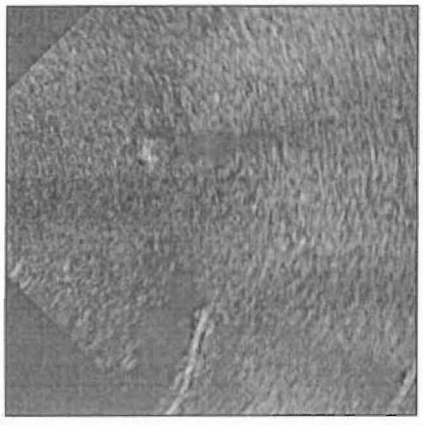
The image quality parameters were evaluated for the original and verifying images. The Image Quality was evaluated using the Image Quality Equation and is shown in Figure 7.4. The original image for all slices is Image 1. The 2^{nd} and 3^{rd} images were designed to be better than the original and Images 4–6 were designed to be worse than



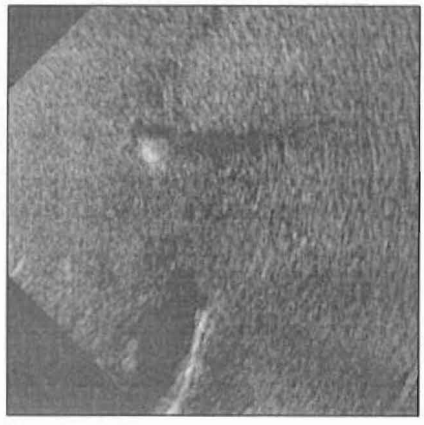
(a) Slice 019 Image 1



(b) Slice 019 Image 2



(c) Slice 019 Image 3

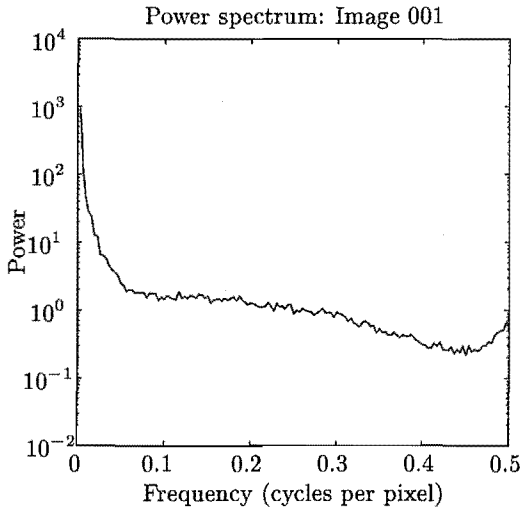


(d) Slice 019 Image 4

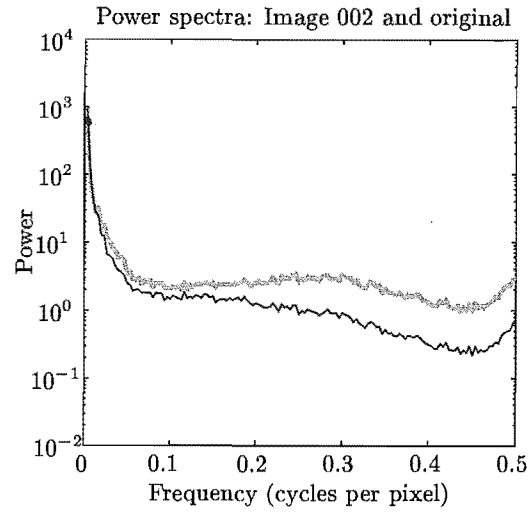
(e) Slice 019 Image 5

(f) Slice 019 Image 6

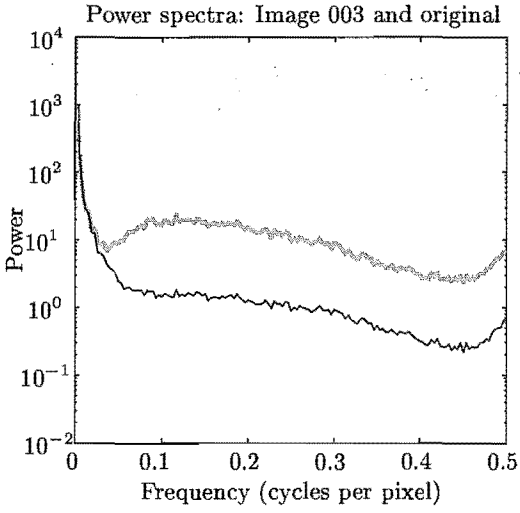
Figure 7.2 Set of ranked images.



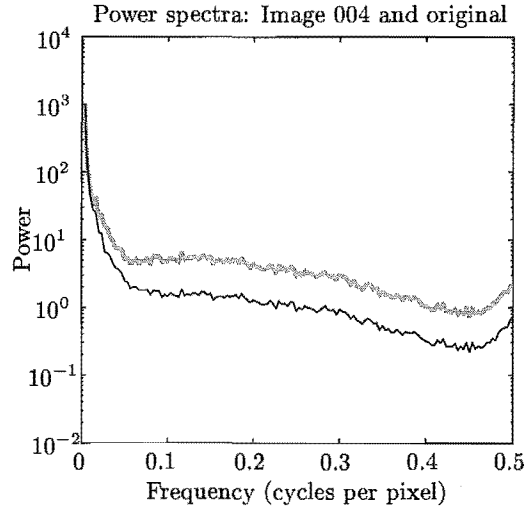
(a) Slice 019 Image 1



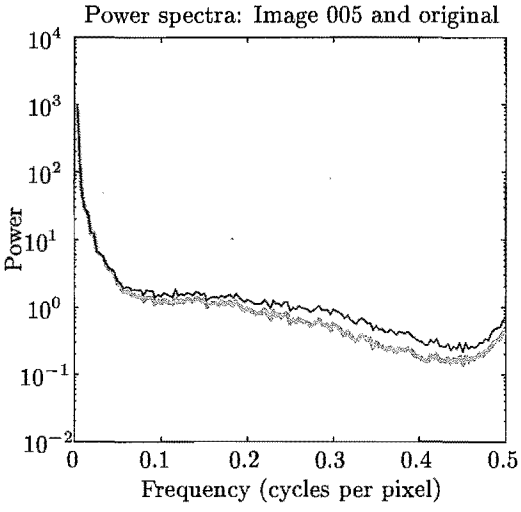
(b) Slice 019 Image 2



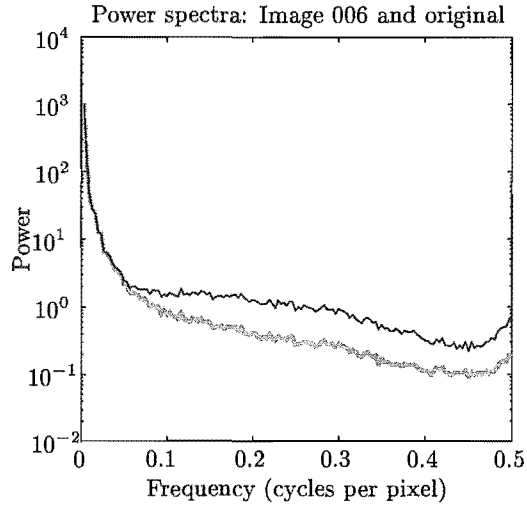
(c) Slice 019 Image 3



(d) Slice 019 Image 4



(e) Slice 019 Image 5



(f) Slice 019 Image 6

Figure 7.3 Power spectra of images in Figure 7.2.

the original according to the assessment criteria. For each image number, the same manipulations were carried out and the resulting characteristics are similar for each slice. One of the slices and accompanying verifying images are shown in Figure 7.5.

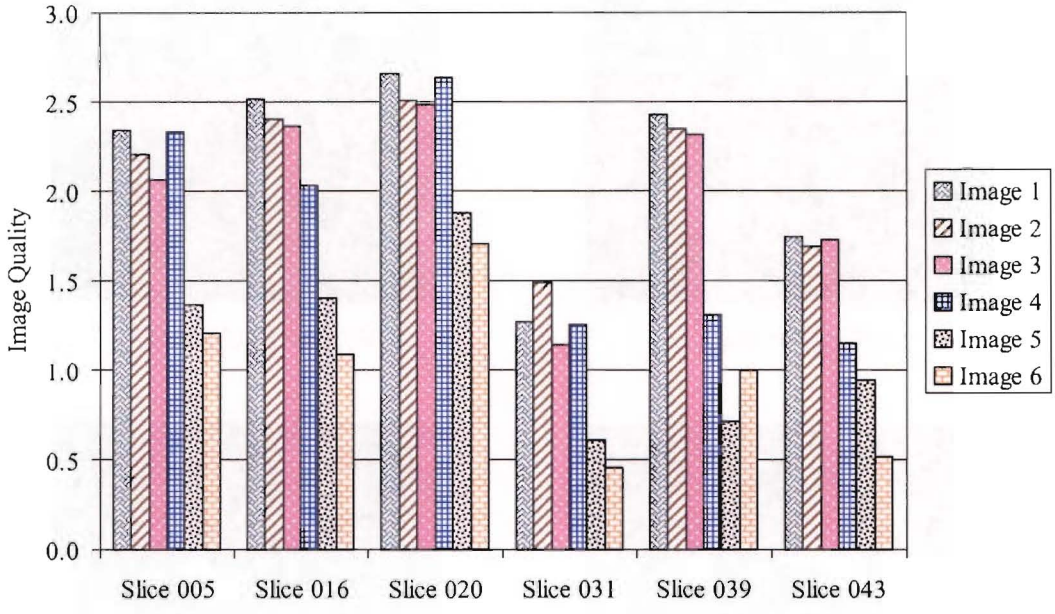


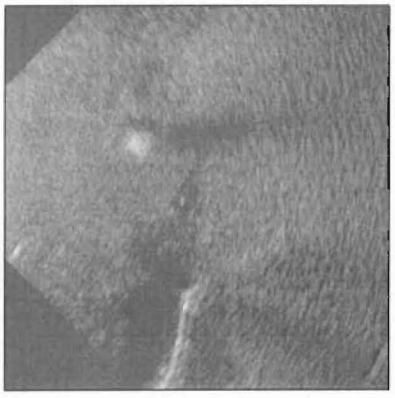
Figure 7.4 Image Quality of verifying images evaluated by the Image Quality Equation. Image 1 is the original image. Images 2 and 3 were designed to be better than the original, and Images 4–6 were designed to be worse than the original.

The Image Quality, as assessed by the IQE, does provide some indication of the quality of the image, despite having a limited range (0–3). A subjective assessment was performed by three observers on a subset of the images, but the process was not as rigorous as the previous image grading. Images 2 and 3, which were subjectively judged to be better than the original, were consistently evaluated as close to or better than the original. The fourth image, however, which has very low contrast in the background and poor contrast of the uterine cavity has similar Image Quality as the original for three of the slices. The remaining images, which were subjectively judged worse than the original, were evaluated as having a lower Image Quality than the original for all the slices. Thus, IQE has some predictive power for assessing the quality of ultrasound images.

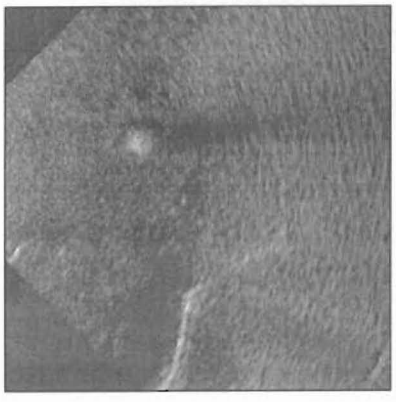
7.2.3 Discussion of Image Quality Equation

The Image Quality Equation proposed in Equation (7.17) was designed to provide an objective measure of image quality. The equation does not have a good statistical fit for predicting image quality. The poor fit is due to high observer variability and independent variables that do not measure image quality accurately. However, the

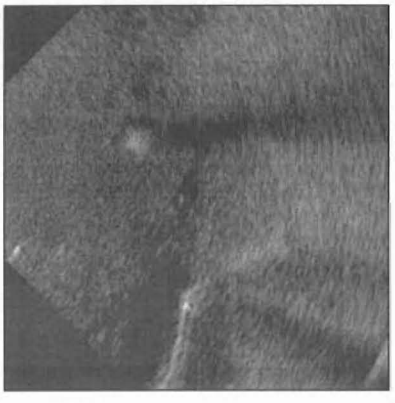
(a) Slice 020 Image 1



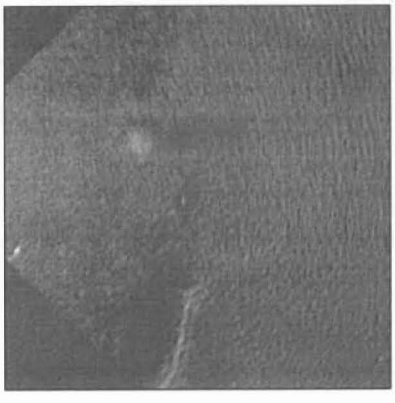
(b) Slice 020 Image 2



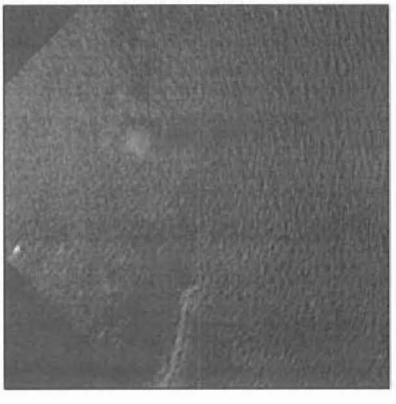
(c) Slice 020 Image 3



(d) Slice 020 Image 4



(e) Slice 020 Image 5



(f) Slice 020 Image 6

Figure 7.5 Original image of the 20^{th} slice and the accompanying verifying images. (a) Original image. (b–c) Enhanced images of the original image. (d–f) Images designed to be worse than the original image.

results using the verifying images showed that images that were subjectively judged better were rated with a higher Image Quality than images that were judged worse for a particular slice. This finding suggests that the Image Quality Equation has some value, but overall, further research is needed.

Procedures that could lead to improvement include increasing the range for the Image Quality, reducing the intra- and inter-personal differences in the grading, and adjusting the equation parameters to provide a better model of image quality. A possible method for achieving a greater range for the Image Quality, reflecting the image quality grades given to the grading images, may be by selecting a different set of graded or ranked images. Currently, both sets of images contain a large proportion of low quality images. A more even spread of image quality in the images may have resulted in a greater range in the Image Quality.

The graded images may have had qualities that varied across the range too widely, for instance the blurred images, which were ranked as 3 and 4, were difficult to compare against images with sharp edges, but with no contrast in the background (ranked 1, worst). The background of the image ranked 2 also looked fairly similar to the background of the image ranked 5 (best), even though the contrast of the features of interest was better. Conflicting characteristics (blurred but good contrast, or sharp but poor contrast) made images difficult to grade. Therefore, a set of images with widely varying image quality, but preferably without conflicting characteristics, would probably prove to be easier to grade consistently.

Choosing people who are familiar with the type of image being graded, or training people better for the task may provide more consistent results, especially if there are not conflicting characteristics in the images.

The investigation into the power spectra of the images revealed that an enhanced image does not necessarily have a lower power than the original image. The investigations also suggested that the shape of the power spectra might be better utilised in the Image Quality Equation.

None of the problems raised in the discussion points are trivial to solve, however, they do indicate further avenues for research with an aim to produce a better objective measure of image quality.

7.3 CONCLUSIONS

An approach towards objectively assessing image quality was presented. The approach was to develop a regression equation based on several measurable features of ultrasound images and rankings provided by independent assessors. Although the approach resulted in only a mediocre objective measure for ultrasound images, several directions for future research in this area have been identified.

A successful image quality equation would be able to be used to assess the performance of various wavelets and the enhancement schemes presented in the following chapter.

Chapter 8

ULTRASOUND IMAGE ENHANCEMENT

This chapter describes the application of two-dimensional wavelet transforms using both separable and non-separable wavelets for the purpose of improving the quality of a range of images. Noise reduction and contrast enhancement functions, collectively termed enhancement techniques, are developed and employed, with particular emphasis on improving ultrasound images. Different categories of wavelets are used to determine whether they influence the performance of the methods. However, the emphasis of the chapter is on developing the enhancement techniques. The new enhancement techniques are compared with existing methods. The applicability of the new methods for improving the quality of ultrasound images, and the advantages and disadvantages of the methods are discussed.

There are four wavelets used in this chapter. They are can be described in four categories: separable linear phase, separable orthogonal, non-separable linear phase and non-separable orthogonal. The separable wavelets are constructed using tensor products of the Daubechies order 2 wavelet and the Biorthogonal 1.3 wavelet, which are orthogonal and linear phase, respectively. These wavelets will be denoted *daub2* and *biorl.3* in the text, but it should be remembered that in this context, the wavelets are two-dimensional wavelets. The non-separable linear phase wavelet, *kovalinp*, is defined in Equation (6.22), and the non-separable orthogonal wavelet, *kovaorth*, is defined in Equation (6.14), with the solution given in Equation (6.16). The separable wavelets were chosen to match the non-separable wavelets in size; the construction of these wavelets was described in Sections 6.1 and 6.2.

In the first section, noise reduction techniques are reviewed and two novel techniques introduced. The methods are tested and compared. In the second section, enhancement techniques are described including two new functions for contrast enhancement. The enhancement techniques are applied to ultrasound images for evaluation in the third section. Finally, conclusions are given in the last section.

8.1 NOISE REDUCTION TECHNIQUES

In this section, existing and novel wavelet based noise reduction techniques are presented. These techniques use characteristics of the wavelet transform, namely, band pass filtering, to narrow the application of noise reduction techniques to distinct spectral bands in the image. These methods rely on the assumption that noise occurs in separate spectral bands from the image data. Five methods of noise reduction are discussed. The first three methods are common in the literature and are described in the first subsection. Two new methods are described in the second subsection. The first is an extension of a one-dimensional scheme presented recently in the literature, and the second is a new application of a common enhancement scheme. The novel methods are tested and compared with one of the existing methods in the third subsection and the results discussed in the fourth subsection.

8.1.1 Existing Noise Reduction Methods

This section describes the existing noise reduction methods.

One method of noise reduction is to set the levels of the wavelet transform associated with noise to zero and enhance the remaining levels [Laine *et al.* 1995, Kang *et al.* 1998]. This approach is based on the assumption that the noise occurs in a different spectral bands than image features. Unfortunately, edge information is often in the same levels as noise. Therefore, discarding these levels entirely results in an unsatisfactory blurring of the image [Laine *et al.* 1995].

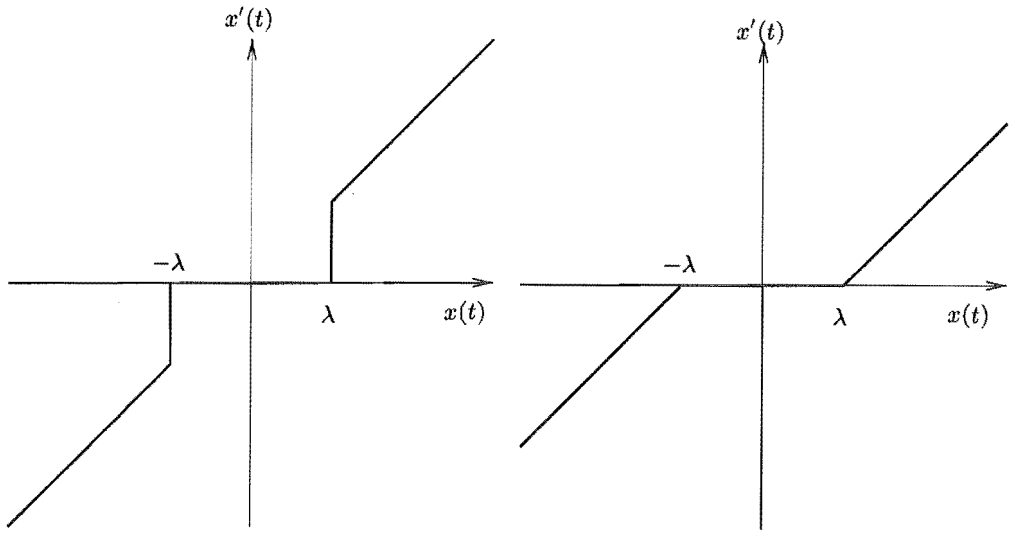
A second method of noise reduction is *hard-thresholding*, which is based on the assumption that the wavelet coefficients due to noise have a smaller magnitude than the coefficients due to features in the image. Hard-thresholding simply reduces to zero any signal component, $x(t)$, which is less than a threshold, λ, whilst keeping the other values. The new values, $x'(t)$, are expressed as

$$x'(t) = \begin{cases} x(t), & \text{if } |x(t)| > \lambda, \\ 0, & \text{otherwise.} \end{cases} \tag{8.1}$$

Donoho [1995] introduced *de-noising* using *soft-thresholding*. Soft-thresholding reduces to zero any signal component that is less than a threshold and reduces all other components towards zero. For a signal, $x(t)$, the soft-threshold function is given by:

$$x'(t) = sign(x(t))max(0, |x(t)| - \lambda), \tag{8.2}$$

where $sign(x(t)) = \pm 1$ according to the sign of $x(t)$, and λ is the threshold value. The two threshold functions are displayed in Figure 8.1.



(a) Hard-thresholding function (b) Soft-thresholding function

Figure 8.1 Thresholding functions.

Donoho [1995] performs soft-thresholding at all levels of the wavelet transform in the wavelet domain. The modified wavelet coefficients are then reconstructed in the standard manner. The property of perfect reconstruction is lost, but for carefully chosen thresholds, the reconstructed image is enhanced.

The threshold value, λ_n, defined by Donoho [1995] is given by:

$$\lambda_n = s_m \sqrt{2\log(n)}, \tag{8.3}$$

where n is the number of coefficients and s_m is given by the variance of the wavelet coefficients $Var(d_{ji}) \leq s_m^2, i = 0, \ldots n$ and for each scale, j. This approach has been used by others successfully [Herrera *et al.* 1999].

Another estimate for the threshold is based on the sample variance, s_s of the lowest scale of the wavelet transform coefficients [Chang and Vetterli 1997]. The sample variance s_s is given by:

$$s_s^2 = \frac{\sum (d[n_1, n_2] - \bar{d})}{N},$$

where

$$\bar{d} = \frac{\sum d[n_1, n_2]}{N} \tag{8.4}$$

and N is the number of samples. The sample variance is used because it is a reasonable estimate of the noise power of the image [Chang and Vetterli 1997]. The estimate of the noise power at other levels can be calculated using the known decomposition filters.

Chang and Vetterli [1997] use the following for uniform thresholding:

$$\lambda = 1.8 s_s. \tag{8.5}$$

Laine *et al.* [1995] and Chang and Vetterli [1997] use Donoho's soft-thresholding technique only in the low scales. The reason is that noise can, in general, be expected to be concentrated only in low scales that correspond to the high frequency part of the transform.

In this thesis, for noise reduction performed using soft-thresholding, the threshold, λ, is determined by Equation (8.5), that is, a constant multiplied by the sample variance of the first level of the wavelet transform coefficients. The constant is determined in Section 8.1.3.

8.1.2 Novel Methods of Noise Reduction

This section describes two novel methods of noise reduction. The first method, Two-Dimensional Derivative Denoising, is an extension of a one-dimensional method reported recently in the literature. The second method, Wavelet Median Denoising, applies median filtering in the wavelet domain. These methods are explained in more detail in the following two subsections.

8.1.2.1 Two-Dimensional Derivative Denoising

The first novel method of noise reduction, *Two-Dimensional Derivative Denoising*, is based on one-dimensional derivative noise reduction [Roy *et al.* 1999]. Denoising a one-dimensional noisy signal is accomplished by modifying the derivative of the signal in the wavelet domain. The modified derivative is reconstructed by the inverse wavelet transform and the result integrated to give the denoised signal.

In the one-dimensional case, the derivative, $x_d(t)$, of a signal $x(t)$ is obtained, $x_d(t) = \frac{dx(t)}{dt}$. The derivative is transformed to the wavelet domain (Equation (2.23)) giving wavelet coefficients X_d. At this point, the power at each scale, or *scalar power*, is estimated at each scale, j, and displacement, τ, by:

$$P_j(x) = \sum_{\tau=-\infty}^{\infty} \|X_{d:j,\tau}\|^2 \quad (j = 1, 2, 3, \ldots) \tag{8.6}$$

The underlying assumption of this method is that the noise and the data in the derivative are more highly concentrated in separate scales of the wavelet domain than in the original signal. This occurs for white noise as differentiation transforms uncorrelated stochastic processes to first order moving average processes. More energy is therefore distributed to the finer scales [Roy *et al.* 1999]. The appropriateness of this

method was demonstrated in one-dimension [Roy *et al.* 1999] by plotting the power of the signal in the wavelet domain at each scale and the power of the derivative of the signal in the wavelet domain at each scale. These two plots were compared with plots done with the same signal with added noise. In the noisy signal, the noise was spread up the scales of the wavelet transform. However, in the derivative of the noisy signal, the noise could be seen to be concentrated at the lower scales in the wavelet domain. The scale, j_n at which the power due to noise drops off rapidly is identified and all wavelet coefficients, $X_{d:j,\tau}$ up to scale j_m are set to zero, that is, $X_{d:j,\tau} = 0$, for $j = 1, 2, \ldots, j_m$. The remaining coefficients are used to reconstruct the denoised derivative, $\hat{x}_d$ (Equation (2.24)). To complete the process, the derivative is integrated to give the denoised signal, $\hat{x} = \int \hat{x}_d dt$.

In the extension of the one-dimensional theory to the two-dimensional case, presented here, the derivative is first obtained by calculating the slope and direction at each pixel. To approximate the derivative, the Sobel operator [Russ 1998] is used to find the horizontal and vertical components of the slope. The masks

$$A_{x_1} = \frac{1}{8}\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \text{ and } B_{x_2} = \frac{1}{8}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

are applied to the image, $\mathbf{x} = (x_1, x_2)$. The resulting derivative in the horizontal direction, $a = A \cdot \mathbf{x}$, and derivative in the vertical direction, $b = B \cdot \mathbf{x}$, are then combined to give the magnitude of the slope, $c = \sqrt{a^2 + b^2}$, and the direction at each pixel is also recorded, $\theta = arctan(b/a)$. The magnitude of the slope is used in the algorithm. The magnitude of the slope is transformed using the wavelet transform, C_j, and noise reduction performed on the scales containing noise as for the one-dimensional case, giving modified wavelet coefficients, $\hat{C}_j$. The transform is reversed and the direction recombined with the new magnitude, $\hat{c}$, to determine the new derivatives in the horizontal, $\hat{a} = \hat{c}cos\theta$, and vertical, $\hat{b} = \hat{c}sin\theta$ directions. A cumulative method (see Appendix A) is used to integrate the new derivatives resulting in a noise reduced image, $\hat{\mathbf{x}}$.

8.1.2.2 Wavelet Median Denoising

A second novel noise reduction method is *Wavelet Median Denoising*. The method uses a classical image enhancement technique, median filtering, but applied in the wavelet domain to reduce noise. Median filtering is a statistical approach in which the pixels in a neighbourhood are ranked according to brightness and the central pixel is replaced by the median value of the ranked pixels [Russ 1998]. A schematic of the process is shown in Figure 8.1.2.2. An image $\mathbf{x}[n]$ is decomposed. The median filter, $M_{m,j}$, of width m is applied in the wavelet domain at scales j and the modified wavelet coefficients are reconstructed to form the denoised image, $\mathbf{x}_M[n]$.

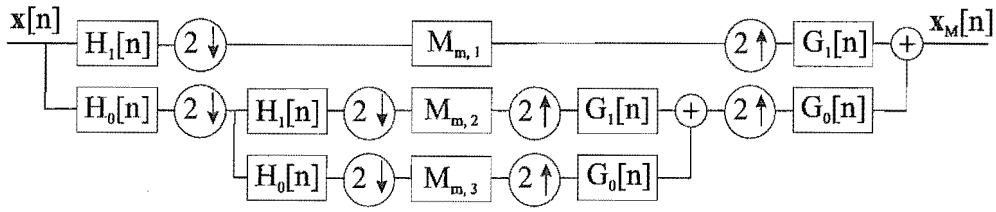


Figure 8.2 Schematic of Wavelet Median Denoising. The image $\mathbf{x}[n]$ is decomposed. The median filter, $M_{m,j}$, of width m is applied in the wavelet domain at scales j and the modified wavelet coefficients are reconstructed to form the denoised image, $\mathbf{x}_M[n]$. $M_{m,j}$ may be different for each scale.

The underlying assumption is that noise causes isolated wavelet coefficients to have high (or low) valued components compared to their neighbours. Conversely, a wavelet coefficient is likely to have adjacent coefficients that have a similar magnitude. This effect is *clustering* in Crouse *et al.* [1998]. For example, wavelet coefficients due to edges will tend to be grouped with several coefficients in the vicinity having a high intensity, and therefore, will tend to have a higher median value. This approach should change the wavelet coefficients due to noise to the same as the local neighbourhood, but should retain the wavelet coefficients due to edges. Therefore, noise in the image will be reduced, whilst the edges of objects will be retained.

The median filter has two parameters that may be varied: the width of the median filter window and the range of scales over which it is applied. A range of values for these parameters is tested in the following section.

8.1.3 Noise Reduction Evaluation

This section presents the results from the noise reduction methods described in the previous sections. Three methods of noise reduction are presented: Two-Dimensional Derivative Denoising, soft-thresholding, and Wavelet Median Denoising. Derivative denoising is a method that was proven to work with a type of one-dimensional signal [Roy *et al.* 1999]. The method was expanded to two-dimensions in Section 8.1.2.1 and the preliminary conditions for success are tested. Soft-threshold denoising [Donoho 1995] has been used successfully in other studies and some results are presented. The results of Wavelet Median Denoising (WMD) are presented. Soft-thresholding and WMD are compared using two test images, a graduated gray scale image and an image of a rose window, and the four wavelets described at the beginning of the chapter. The Signal to Noise Ratio (SNR) is used to compare results. A comparison of the results from both methods is discussed at the end of the section.

8.1.3.1 Two-Dimensional Derivative Denoising

Derivative denoising is performed by calculating the derivative of the image, taking the wavelet transform of the derivative, and reducing the wavelet coefficients in the scales

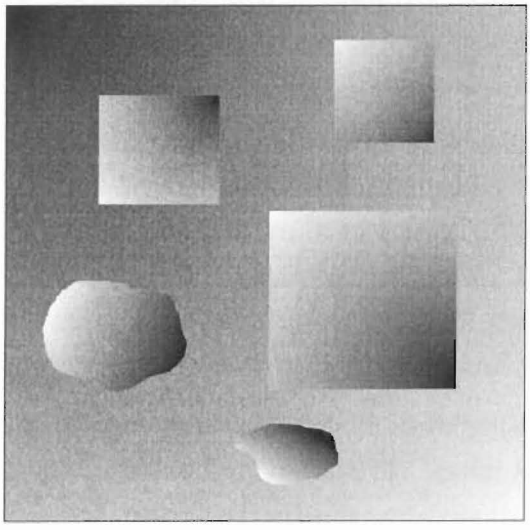
where the noise is concentrated; the inverse wavelet transform is then calculated, and finally, integrated to produce the denoised image (see Section 8.1.2.1). This process is an extension of the one-dimensional algorithm introduced by Roy *et al.* [1999].

The method relies on the assumption that noise in the image is separable from the image objects in the wavelet domain. This assumption is tested by evaluating the power at each scale for an original image and a noisy version of the same image, and the derivatives of the two images [Roy *et al.* 1999]. For the method to be successful, the power due to the noise must be separated from the power due to the objects by more scales in the derivative of the images than in the original images.

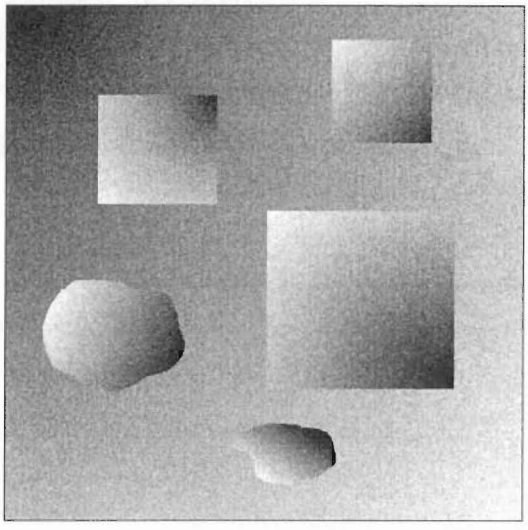
Two different images were used to test the separation of the power due to the noise and objects in the wavelet domain. The first image had a graduated gray scale in the background and five objects with reversed graduated gray scale in the foreground, of size 512x512 pixels, shown in Figure 8.3(a). The second image was the rose window image used in Section 6.3.1, and shown again in Figure 8.3(c). Each image had gaussian noise (mean = 0, variance = 0.005) added to produce the noisy images. The corresponding noisy images are shown in Figure 8.3(b) and 8.3(d).

The four two-dimensional wavelets described at the beginning of the chapter were tested. The tests revealed that for these images and wavelets, the derivative denoising method does not provide any advantages over simple denoising. The power due to the noise was concentrated at low scales for most wavelets for the original images. The power due to the objects was concentrated in the high scales for the kovaorth wavelet, the middle range of scales for the daub2 and bior1.3 wavelets, and the low scales for the kovalinp wavelet. However, for the derivative of the images, the power due to the noise and the objects is concentrated in the lower scales for all the wavelets, thus violating the assumption upon which the derivative denoising method is based. The scales that the power of the noise and the objects were concentrated for the original images and their derivatives are summarised in Table 8.1 for each of the wavelets. The overlap of scales for the peaks in the power due to noise and objects in the image can be seen for all but the separable linear phase wavelet operating on the graduated image. The peaks for the noise and the objects in the bior1.3 wavelet case are, however, closer together in the derivative than the original. An example is shown of the scalar power for the rose window image and the bior1.3 wavelet in Figure 8.4. These findings show the essential preliminary condition for the success of this process was not met.

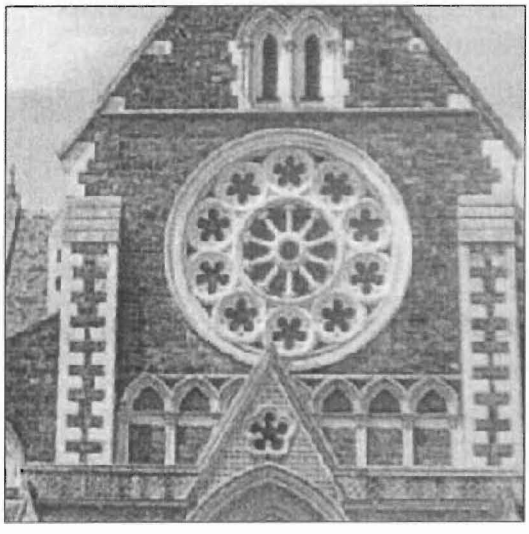
The condition for the derivative denoising scheme to succeed was not met for the two test images. However, ultrasound images do have different characteristics, so the scalar power was also evaluated for an ultrasound image and corresponding derivative. Since this image was already noisy, there was no an original to compare it against. However, the condition for success was still the same—peaks in the scalar power that were more distinct for the derivative than the original. The results are summarised in



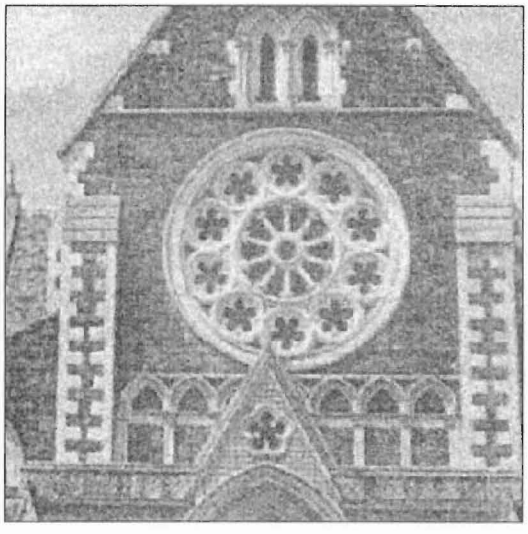
(a) Original graduated gray scale image



(b) Noisy graduating gray scale image



(c) Original rose window image



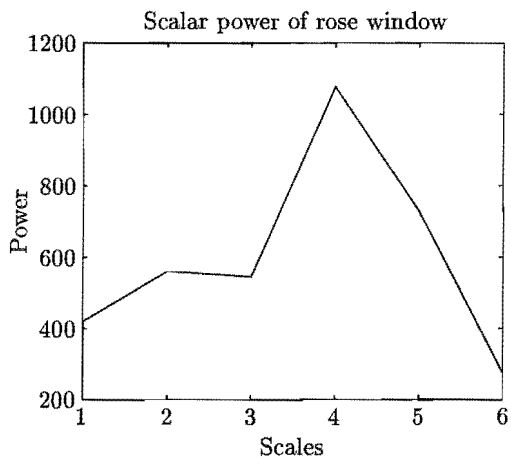
(d) Noisy rose window image

Figure 8.3 Four images used to test the condition for success of two-Dimensional Derivative Denoising.

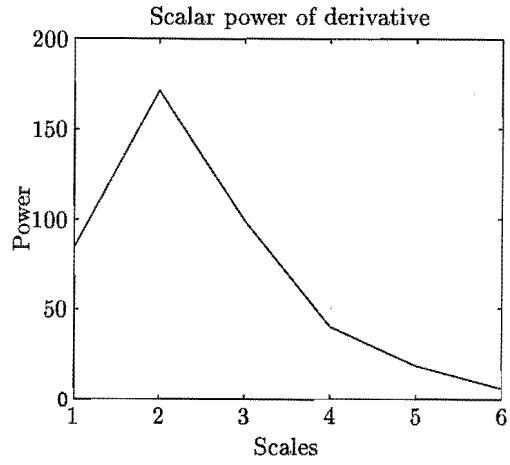
Table 8.2, and again show that the condition for success is not met as either there are not two distinct peaks, or the peaks move closer together. This method was not tested any further.

8.1.3.2 Comparison between Soft Threshold and Wavelet Median Denoising

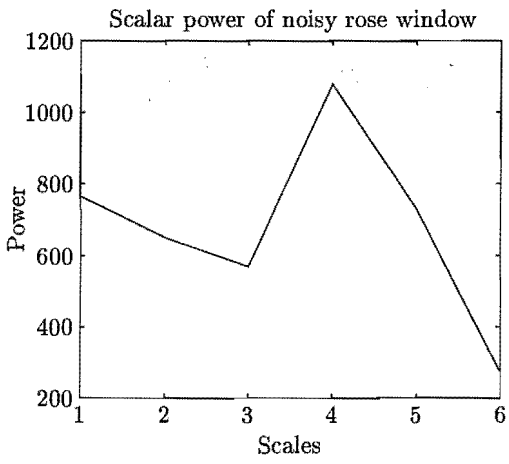
Two methods of denoising are compared in this section. The first is soft-threshold denoising [Donoho 1995] and the second is a novel method, Wavelet Median Denoising



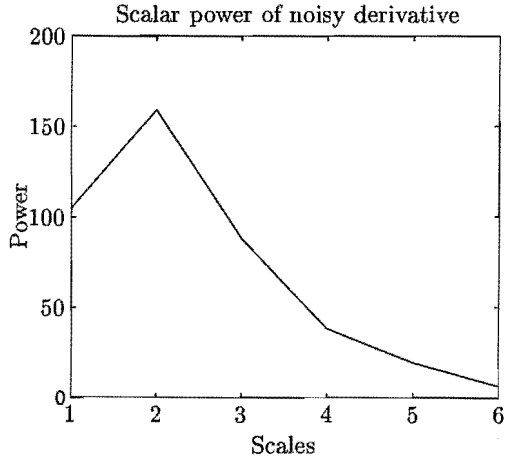
(a) Power at each scale in the wavelet domain of rose window image using separable, linear phase wavelet



(b) Power at each scale in the wavelet domain of rose window image derivative using separable, linear phase wavelet



(c) Power at each scale in the wavelet domain of rose window image with noise using separable, linear phase wavelet



(d) Power at each scale in the wavelet domain of rose window image with noise derivative using separable, linear phase wavelet

Figure 8.4 Scalar power of rose window image (top) with gaussian noise (bottom) analysed using separable, linear phase wavelets

(WMD). The measures are tested subjectively and using the Signal to Noise Ratio (SNR). The methods are executed using the four wavelets under review and, therefore, a comparison between the wavelets is also performed. Both methods are tested using the graduated gray scale image and the rose window image, each with additive gaussian noise, as shown in Figure 8.3. The two images have different characteristics, the rose window image having more high frequency components than the graduated gray scale image.

The four wavelets described at the beginning of the chapter—daub2, bior1.3, kovaorth and kovalinp—are used in both denoising methods. Recall that tensor products

	Graduated image		Graduated derivative	
Wavelet	Noise scales	Object scales	Noise scales	Object scales
Separable, orthogonal	1	4–6	1–2	1–4
Separable, linear phase	1–2	5–6	1	2–3
Non-separable, orthogonal	1–3	8–12	1–2	1–8
Non-separable, linear phase	1	1–2	1	1–2

	Rose window image		Rose window derivative	
Wavelet	Noise scales	Object scales	Noise scales	Object scales
Separable, orthogonal	1–2	2–4	1	1–4
Separable, linear phase	1–2	4–5	1	1–3
Non-separable, orthogonal	1–2	8–12	1	1–6
Non-separable, linear phase	1	1–2	1	1–2

Table 8.1 Range of scales that power is concentrated in the wavelet domain for graduated gray scale image (top), rose window image (bottom), their noisy counterparts, and their derivatives for four wavelets.

	Ultrasound image	Derivative image
Wavelet	Noise scales	Object scales
Separable, orthogonal	2, 5–6	1–3, 6
Separable, linear phase	2, 5–6	1–3, 6
Non-separable, orthogonal	10–12	1–4
Non-separable, linear phase	1–2	1–2

Table 8.2 Range of scales that power is concentrated in the wavelet domain for ultrasound image and derivative for four wavelets.

of the daub2 and bior1.3 wavelets are used to form two-dimensional separable wavelets and that the kovaorth and kovalinp wavelets are non-separable. Both methods are tested for scales up to three decomposition levels for the separable wavelets and up to six decomposition levels for the non-separable wavelets.

Soft-threshold denoising reduces values below a threshold to zero and the values above the threshold towards zero. The method requires a threshold to be set. As discussed in Section 8.1.1, the threshold is the product of a constant and the sample variance of the first level of the wavelet coefficients. Chang and Vetterli [1997] found a constant of 1.8 to be suitable. Here, six constants were tested in a range from 0.2 to 3.

Wavelet Median Denoising is based on classical median filtering. In classical median filtering, for each pixel in an image, a neighbourhood of pixels is ranked according to intensity and the median value replaces the central pixel [Russ 1998]. The classical approach is modified so that the filtering occurs in the wavelet domain at selected scales of a wavelet decomposition. This method is tested for three different neighbourhoods. All neighbourhoods are square, with sizes of 3x3, 5x5 and 7x7 pixels.

In the analysis, it was found that denoising to a low number of scales produced images of better image quality. Images that had been denoised to a high number of scales had a significant amount of blurring. The amount of blurring was proportional to the number of scales denoised. For a given scale, soft-threshold denoising and Wavelet

Median Denoising produced similar results.

The Signal to Noise Ratio (SNR) was used to assess the quality of the denoising. Comparison of the results of denoising between the two images was performed using a Figure of Merit (FOM) [Haykin 1983], given by:

$$FOM = \frac{SNR_d}{SNR_n} \tag{8.7}$$

where SNR_d is the SNR for the denoised image and SNR_n is the SNR for the noisy image. The FOM is greater than one if the denoised image has a better SNR than the noisy image.

Assessing the two methods using the FOM showed that the new method, Wavelet Median Denoising, performed at least as well as soft-thresholding in the lower scales.

The results are shown in Figures 8.5 to 8.8. These figures show the FOM calculated for each level of denoising for each wavelet and median filter window size or threshold weight, as applicable. Figure 8.5 shows the FOM for the graduated gray scale image using Wavelet Median Denoising. Figure 8.6 illustrates the results for the graduated gray scale image using soft threshold denoising. The results for the rose window image are shown in Figure 8.7 for the Wavelet Median Denoising method and in Figure 8.8 for soft threshold denoising.

These results show several interesting features. Firstly, the window size did not make a significant difference in SNR improvement for the Wavelet Median Denoising. Secondly, the different threshold weightings made more of a difference for soft thresholding. In the graduated gray scale analysis, the best results were obtained using a small weighting. However, in the rose window analysis at low scales, changing the threshold weightings made less of a difference to the results. Both non-separable wavelets showed similar patterns over the scales for a given test image and method of denoising. All the wavelets showed a similar response for the rose window image denoised using soft-thresholding. Applying Wavelet Median Denoising to the rose window image gave rise to all four wavelets showing a similar response. However, the responses for the rose window image were different for the two methods. The soft thresholding method produced increasing FOM as the number of scales was increased, while the Wavelet Median Denoising produced decreasing FOM as the number of scales increased. The graduated gray scale image shows similar trends for all four wavelets of a fairly constant FOM over the scales, but with a slight dip in the middle of the number of scales being denoised, except the separable linear phase wavelet which showed a rise.

Although the FOM increased for the soft denoising of the rose window image as the number of scales increased, this is not an indication of better image quality, as the images became progressively more blurred as the number of scales denoised increased.

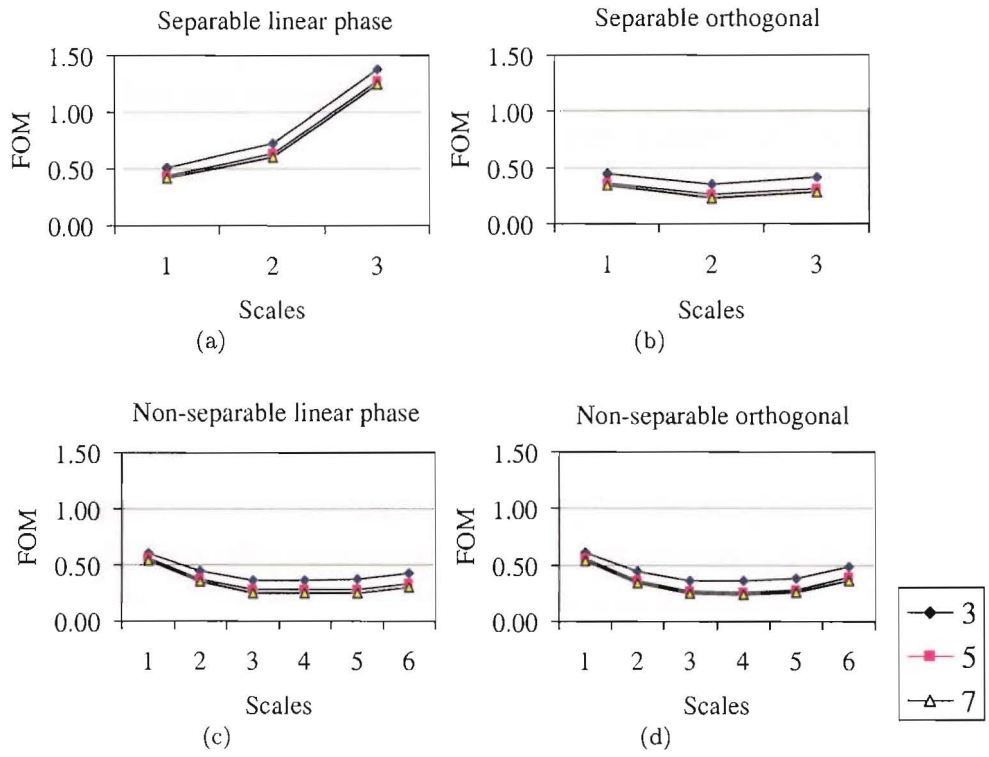


Figure 8.5 FOM for graduated gray scale image denoised using Wavelet Median Denoising.

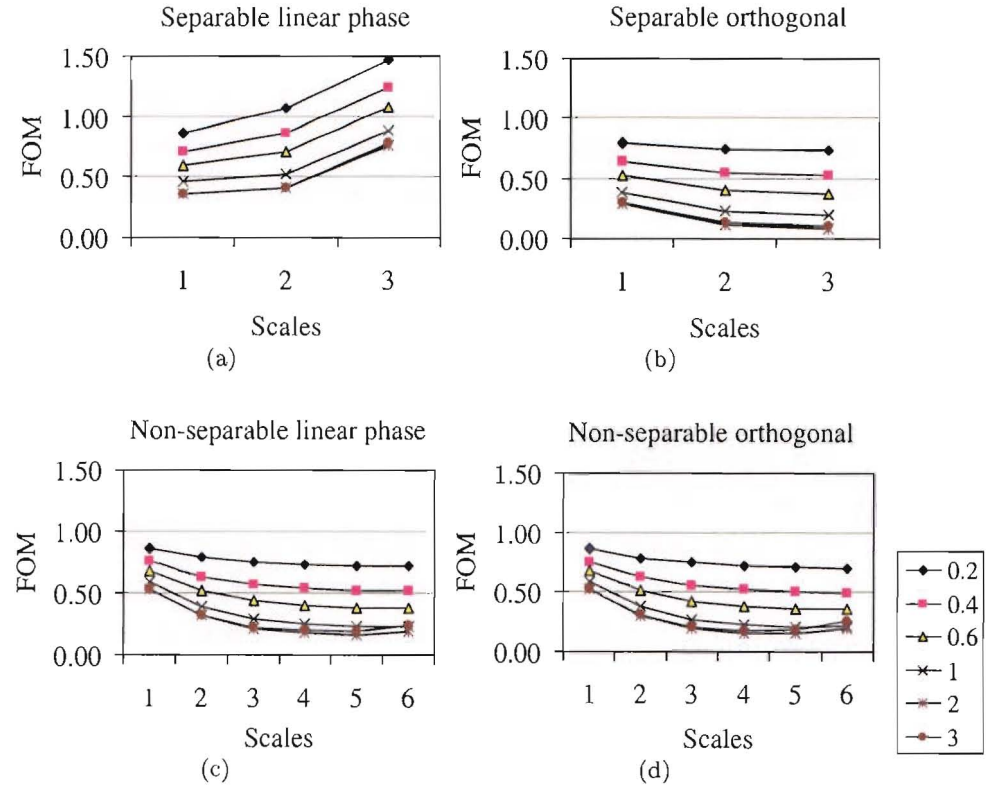


Figure 8.6 FOM for graduated gray scale image denoised using soft threshold denoising.

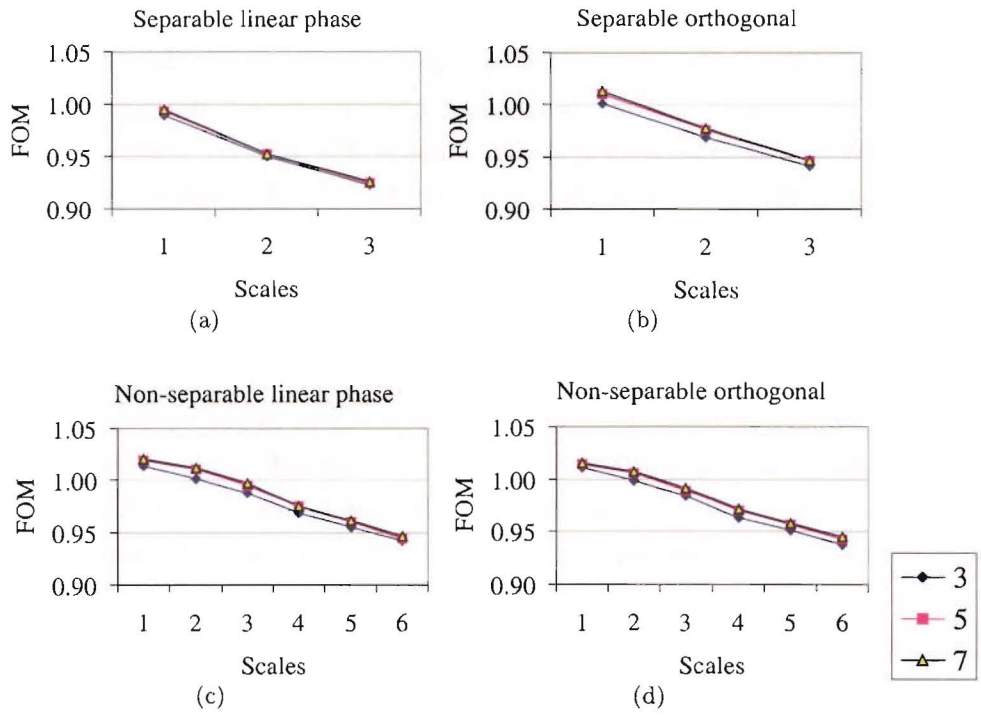


Figure 8.7 FOM for rose window image denoised using Wavelet Median Denoising.

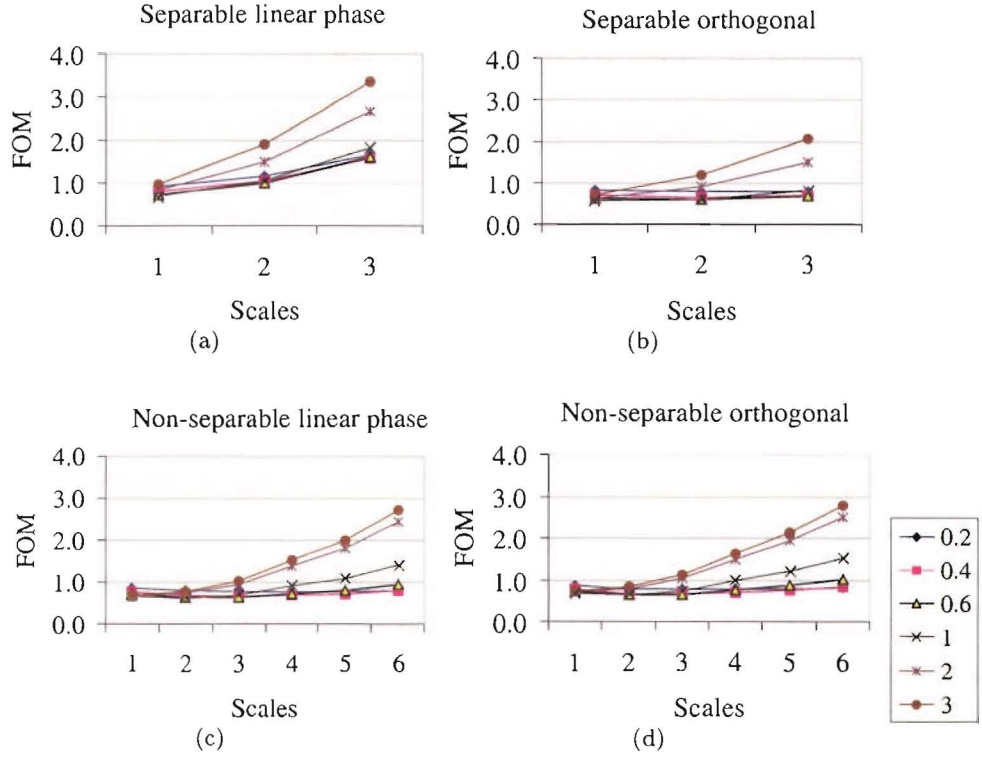


Figure 8.8 FOM for rose window image denoised using soft threshold denoising.

Restricting the discussion to the lower scales, since, as shown in Section 8.1.3.1, the noise is concentrated in the lower scales, for the graduated gray scale image denoised with soft thresholding, a smaller weight produces a better result than larger threshold weights. The Wavelet Median Denoising FOM falls within the range of the soft threshold FOM for the graduated gray scale image. For the rose image, the Wavelet Median Denoising FOM is better than the FOM for the soft thresholding for all but the separable linear phase wavelet.

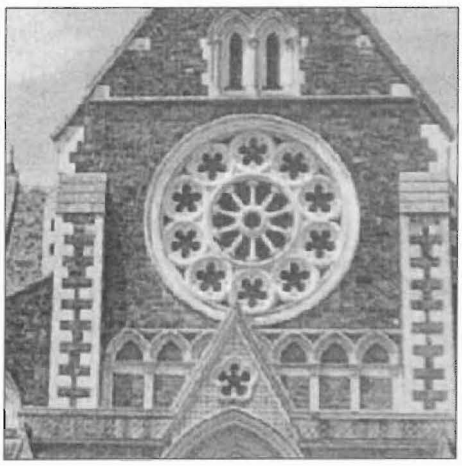
An example of the denoising using the non-separable linear phase wavelet is shown in Figure 8.9. The figure shows the original image and the noisy image at the top for comparison purposes. The middle row shows images denoised using soft-thresholding with a threshold weighting of 3. The bottom row shows images denoised using the Wavelet Median Denoising method with a neighbourhood window size of 3. Figures 8.9(c) and 8.9(e) were denoised to one scale and Figures 8.9(d) and 8.9(f) were denoised to three scales. The images denoised to three scales are blurred; the image obtained by soft-denoising shows greater blurring. The images denoised to one scale have lost some detail compared with the original, but they appear better than the noisy image, with less speckle in uniform areas such as the sky. Both methods seem to have reduced the contrast in the image, with areas that were black in the original and even the noisy image becoming dark gray. The bright areas in the original images seem to have reduced intensity in the reconstructed images.

8.1.4 Discussion of Noise Reduction Methods

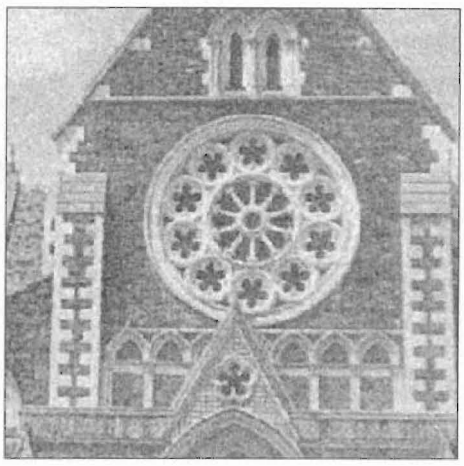
Three methods of noise reduction were investigated: Two-Dimensional Derivative Denoising, soft-thresholding and Wavelet Median Denoising. The methods were tested using two test images with additive gaussian noise and evaluated using the Signal to Noise Ratio (SNR). The SNR is not ideal as it has been shown to be a poor indicator of image quality [Taylor *et al.* 1997], but it does give some indication of the level of noise in an image.

The methods were tested using four two-dimensional wavelets; daub2, bior1.3, kovaorth and kovalinp. There were no significant differences between the performances of the wavelets, aside from the number of decomposition scales for separable wavelets and non-separable wavelets.

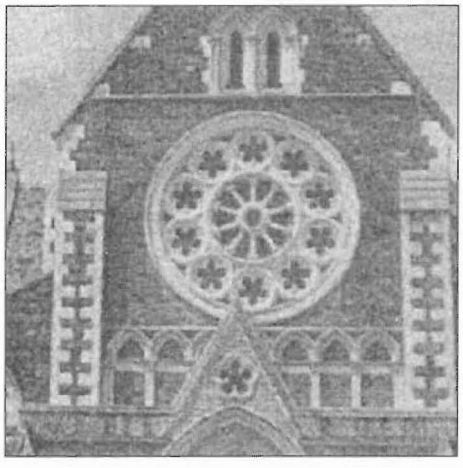
Two-Dimensional Derivative Denoising was tested for a necessary condition for success. The condition required that peaks in the scalar power due to noise and the objects in the image were separated further for the derivative of the signal than for the signal [Roy *et al.* 1999]. However, for the images tested, the peaks converged. This suggests that the noise and the objects did not have sufficiently different ranges of frequency components for the technique to be a success. This may have been compounded by the relatively small size of the images, meaning that the image could only



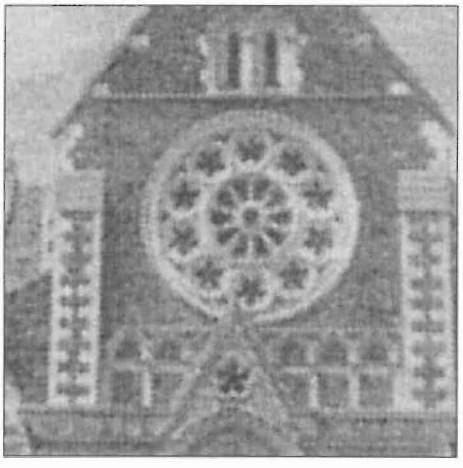
(a) Original rose window image



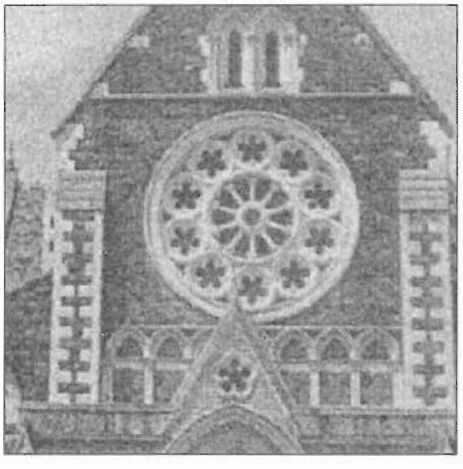
(b) Noisy rose window image



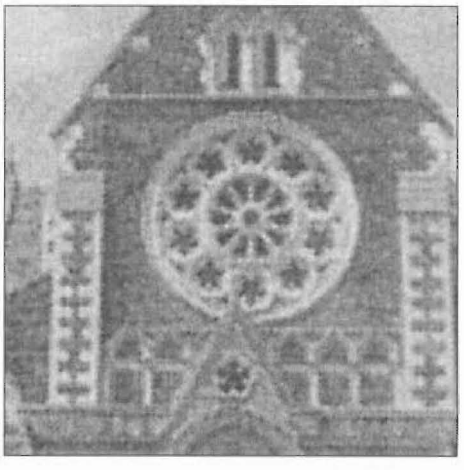
(c) Noisy image denoised using soft-thresholding, threshold weight = 3, 1 scale



(d) Noisy image denoised using soft-thresholding, threshold weight = 3, 3 scales



(e) Noisy image denoised using WMD, window of 3, 1 scale



(f) Noisy image denoised using WMD, window of 3, 3 scales

Figure 8.9 Comparison of soft threshold denoising and Wavelet Median Denoising (WMD).

be decomposed a small number of times. This problem is especially compounded for the separable wavelets. A higher resolution of the images might improve the distinction between peaks for the derivative. Although this technique was not suitable for the images tested here, for high resolution, large size images the technique may be suitable. Another possible explanation for the poor result is that the Sobel operator did not provide a reliable estimate of the derivative. Since the image is noisy and features are blurred, more sophisticated methods to estimate the derivatives may produce improved results. A more sophisticated integration technique may also be required. However, employing more sophisticated methods for approximating the derivative and integral implies loosing the simplicity of the method.

The soft-thresholding method was tested for a range of threshold weightings and scales. Denoising of a large number of scales produces a blurred image. This is to be expected as each wavelet coefficient at higher scales corresponds to a greater area in the image than a wavelet coefficient at lower scales.

A novel method for noise reduction, Wavelet Median Denoising, was tested for a range of neighbourhoods and scales. Again, for images denoised over a large number of scales, the reconstructed image was blurred. Wavelet Median Denoising performed comparably with soft thresholding. However, neither method was outstanding for the images tested here.

The denoised images for both methods appear to have lost contrast. This occurs because the denoising operation causes the range of intensity values of the reconstructed images to be increased at both ends from the range of intensity values for the original images. Therefore, an area that may have been at the bottom of the range of intensity values and appeared black in the original image, may now be elevated in the range and appear dark gray. A similar operation can occur at the high end of intensity values, causing the brightest values to effectively lose intensity.

A possible improvement for both the soft-thresholding denoising and Wavelet Median Denoising methods would be to provide a qualifier for the denoising operation. For a particular pixel in the image, if the pixel showed evidence of being part of an edge, then the pixel would be unchanged. However, if the pixel was not part of an edge then the denoising operation would be performed. Chang and Vetterli [1997] have used adaptive thresholds for soft thresholding, based on the texture of the area in the image. A similar adaptive measure could be used in the Wavelet Median Denoising, adjusting the neighbourhood size for the median filtering, or adjusting the percentile value chosen as the replacement pixel depending on the local texture.

As well as methods estimating the texture of the area in the image, other more recent techniques involve advanced statistical analysis of the wavelet coefficients. In one approach, non-parametric statistical models are used to estimate complexity priors which are applied as a penalty function in the wavelet domain [Liu and Moulin 1997,

Moulin and Liu 1999]. Other techniques use hierarchical statistical analysis of the surrounding pixels in the wavelet domain applying scale-space mixture modeling [Liu and Moulin 1999] and of both surrounding pixels and pixels in adjacent layers of the wavelet domain applying Hidden Markov Models to the analysis [Crouse *et al.* 1998]. These methods may provide even better denoising, but with an associated increase in computational expense.

8.2 CONTRAST ENHANCEMENT TECHNIQUES

Contrast enhancement techniques are designed to enhance the contrast of objects in the image relative to the background. In this section, a method of contrast enhancement developed by Laine *et al.* [1995] is described, which manipulates the wavelet coefficients from x-ray mammography images. A new function for contrast enhancement is developed by adapting Laine's enhancement function, producing a function of enhancement that is suitable for manipulating the wavelet coefficients from ultrasound images. A second new function for contrast enhancement is then described which is designed to reduce the effect of artefacts found in ultrasound images that vary with distance from the ultrasound probe. This function is a spatially varying contrast enhancement technique. The contrast enhancement technique with the novel functions are evaluated.

8.2.1 Existing Contrast Enhancement Technique

Laine *et al.* [1995] used Donoho's soft-thresholding technique in conjunction with contrast enhancement techniques for separable wavelets to enhance x-ray mammograms. Their guidelines for designing an enhancement function were:

- An area of low contrast should be enhanced more than an area of high contrast.
- A sharp edge should not be blurred.

Laine also imposed the following constraints:

- Monotonicity, in order not to change the position of local extrema, nor create new extrema.
- Anti-symmetry, $E(-t) = -E(t)$, in order to preserve phase polarity for "edge crispening."

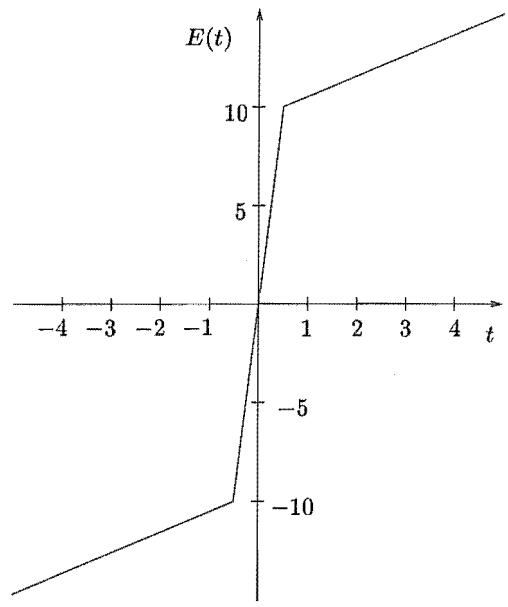


Figure 8.10 Enhancement function used by Laine with $T = 0.5$ and $\mathbf{K} = 20$.

They proposed a simple piecewise non-linear enhancement function, E(t), with threshold value, T, such that

$$E(t) = \begin{cases} t - (\mathbf{K}-1)T, & \text{if } t < -T, \\ \mathbf{K}t, & \text{if } |t| \leq T, \\ t + (\mathbf{K}-1)T, & \text{if } t > T. \end{cases} \tag{8.8}$$

where the gain, $\mathbf{K} > 1$. The enhancement function is shown in Figure 8.10.

In practice, a constant gain, $\mathbf{K}$, was used over all levels, with the threshold, T, adjusted for each level. The threshold T was set so that a constant percentage of the wavelet coefficients would be above the threshold for each level. This percentage and the gain $\mathbf{K}$ were user defined.

Noise reduction is used in conjunction with the contrast enhancement serially. That is, the wavelet coefficients are noise reduced then contrast enhanced, as is shown in Figure 8.11.

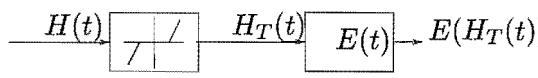


Figure 8.11 Block diagram for noise reduction followed by enhancement.

Laine's approach was considered for ultrasound images. However, two limitations were found. The first limitation was that Laine's first aim to enhance low contrast material and reduce high contrast material does not apply to the ultrasound images analysed in this thesis. The second limitation of Laine's function, specific to ultrasound

images, is that it does not take into account the problem of beam widening. These limitations are discussed in the following sections.

8.2.2 Novel Functions for Contrast Enhancement

In the previous section, a contrast enhancement technique [Laine *et al.* 1995] designed for enhancing mammography images was introduced. There were two limitations noted with the technique with regards to applying it to ultrasound images. The first limitation was the emphasis of the contrast enhancement. The second is related to specific artefacts found in ultrasound images. These two limitations are discussed and alternative functions proposed specifically for the enhancement of ultrasound images.

8.2.2.1 Contrast Enhancement Function

This section describes a method that overcomes the first limitation of Laine's method. The first limitation is due to the nature of the ultrasound images that are tested in this thesis. Laine's method was applied to digital mammograms which have a high intensity background compared with the objects of interest and therefore, a low contrast. Ultrasound images have a background that has a high level of noise, but is essentially of low intensity. Therefore, the contrast between the objects of interest and the background is high compared to the mammogram images. The problem with ultrasound images is to enhance the objects of interest by increasing, or at least maintaining the intensity of those objects, whilst decreasing the intensity of any noise or artefacts in the image.

A non-linear function is proposed where the wavelet coefficients below a threshold are decreased relative to the wavelet coefficients above the threshold. This function retains the constraint of monotonicity and preserves edges. The proposed function, $E(t)$, is given by:

$$E(t) = \begin{cases} \mathbf{K}_1 t, & \text{if } \|t\| < T, \\ \mathbf{K}_2 t, & \text{if } \|t\| > T \quad \text{with } 0 < \mathbf{K}_1 < \mathbf{K}_2. \end{cases} \tag{8.9}$$

By choosing $\mathbf{K}_1$ less than $\mathbf{K}_2$, it is possible to decrease the wavelet coefficients associated with noise and artefacts relative to the wavelet coefficients associated with the objects of interest. This manipulation of the wavelet coefficients is performed for each scale, resulting in an enhanced reconstructed image. The new enhancement function is shown in Figure 8.12.

The threshold and gain values were determined experimentally. The threshold values were set using statistical measures. For each scale, a linear combination of the mean and standard deviation of the transform at that scale was used to determine the threshold. The linear combination was determined by examining the histograms of the transforms and ensuring an appropriate proportion of the wavelet coefficients would

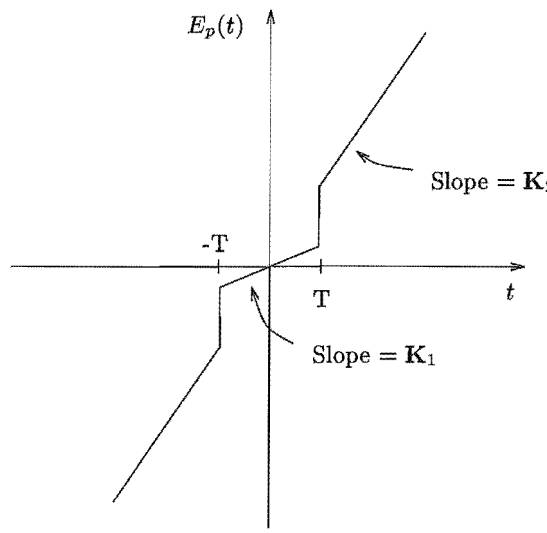


Figure 8.12 Enhancement function used for ultrasound images.

be increased. For each scale, the percentage of wavelet coefficients to be increased depended on the appearance of the objects of interest in the wavelet domain at that scale. A relatively small percentage of the wavelet coefficients were increased at the scales where the objects of interest were not apparent. A relatively high percentage of the coefficients were increased at the scales where the objects of interest were apparent. This function is tested in Section 8.3.2.

8.2.2.2 Spatially Varying Contrast Enhancement

In this section, a new function is proposed for spatially varying contrast enhancement in the wavelet domain. This function is developed to reduce the effect of beam widening in ultrasound images. Beam widening is an artefact that varies linearly with distance from the focal zone of the image. Therefore, a contrast enhancement method that varies spatially is appropriate for reducing this artefact.

Chang and Vetterli [1997] spatially adapt the noise reduction threshold depending on classification of the image into smooth, texture and edge regions. The reasons for spatially varying the threshold are that noise is visually more disturbing in smooth regions, whilst over-smoothing of edges is also visually disturbing. Therefore, the threshold for the smooth regions is made larger, the threshold for the edge regions is made smaller, and the threshold for the textured regions is kept between these two levels.

Ultrasound images have a high level of noise and artefacts. They are difficult to segment into the regions described by Chang and Vetterli [1997]. However, the artefact of beam widening is dependent on spatial location in the image. As the beam travels further away from the focal zone, it gets wider and tends to increase the intensity of the image outside the focal zone. Therefore, a spatially varying enhancement function is proposed.

The new enhancement function is based on the position in the image compared to the position of the focal zone. The focal zone is defined as the area of the image which is in focus. This can be varied using different settings for the scanner resulting in different areas of the image being in focus. This technique assumes the focal zone is positioned close to the scanner. The distance in number of pixels away from the focal zone is denoted by i. Two parameters are varied with distance in the new function. Firstly, as the distance away from the focal zone increases, the threshold value, T_i, increases. Secondly, as this distance increases, the gain, $\mathbf{K}_{s_i}$, is increased. The new function, $E_s(t)$, is given by:

$$E_s(t) = \begin{cases} \mathbf{K}_{s_i}^{-1} t, & \text{if } t < T_i, \\ \mathbf{K}_2 t, & \text{if } t > Ti, \end{cases} \tag{8.10}$$

where

$$T_i = T_{i-1}\mathbf{k}, \tag{8.11}$$

$$\mathbf{K}_{s_i} = \mathbf{K}_{s_{i-1}}\mathbf{k}, \tag{8.12}$$

with $\mathbf{K}_{s_i} > 1$ and $\mathbf{k} > 1$. The value $\mathbf{k}$ is set so that the intensity of the background outside the focal zone is reduced to the same level as the general background within the focal zone. Inside the focal zone, $i = 0$ and the values for the gain and threshold are given by Equation (8.9), that is, $\mathbf{K}_{s_0}^{-1} = \mathbf{K}_1$ and $T_0 = T$. Therefore, inside the focal zone, the spatially varying contrast enhancement function is unchanged from the previous enhancement function. Outside the focal zone, as the intensity of the beam widening tends to increase with distance away from the focal zone, the threshold is increased and the wavelet coefficients below the threshold are decreased by a larger amount as the distance away from the focal zone is increased. This function is tested and evaluated in Section 8.3.2.

Three functions for contrast enhancement have been presented. These functions are compared in the following section for improving image quality.

8.3 IMAGE ENHANCEMENT OF ULTRASOUND IMAGES

In this section, a comparison of four different wavelets is performed using the new contrast enhancement function presented in Section 8.2.2.1. The four wavelets are described at the beginning of the chapter and can be categorised by: separable linear phase, separable orthogonal, non-separable linear phase, and non-separable orthogonal wavelets. The graduated gray scale image is used for the initial comparison.

Three functions of image enhancement presented in Section 8.2 are tested. The method introduced by Laine *et al.* [1995] is tested using the original contrast en-

hancement function against the new contrast enhancement function presented in Section 8.2.2.1 and the new function, spatially varying contrast enhancement, presented in Section 8.2.2.2. The non-separable orthogonal wavelet and an ultrasound image are used for comparison between the enhancement functions.

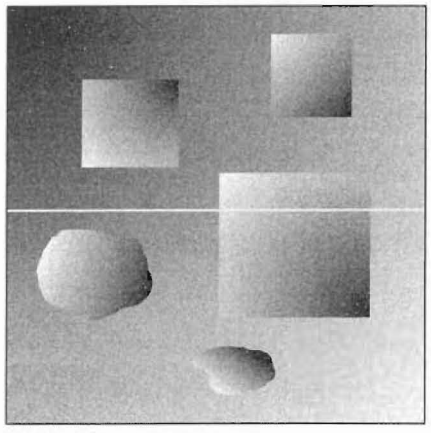
Spatially varying contrast enhancement is tested on different ultrasound images for a limited range of parameters.

8.3.1 Comparison of Different Wavelets

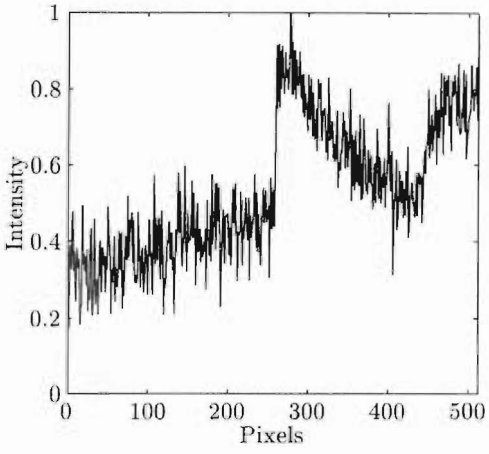
A comparison of the four two-dimensional wavelets, daub2, biorl.3, kovaorth and kovalinp, described at the beginning of the chapter using the new contrast enhancement method was carried out. The noisy graduated gray scale image (512x512), shown in Figure 8.13(a), was used as a test image. A set of parameters were chosen for the enhancement procedure and these were applied using all four wavelets.

An initial study using a small range of parameters, based on the findings in Section 6.3.1, was carried out using the linear phase wavelets. The kovalinp wavelet had a greater range of possible settings allowing the enhancement to be more highly customised. A set of parameters was determined that improved the test image for the kovalinp wavelet. These parameters were then used for all four wavelets. It was found that between the non-separable wavelets there was no significance difference in performance, although the kovaorth wavelet reduced the noise slightly more than the kovalinp wavelet. An original noisy image and its profile are shown in Figures 8.13(a) and 8.13(b). The results for the non-separable wavelets are shown in Figures 8.13(c) and 8.13(e). Profiles of the two enhanced images are shown in Figures 8.13(d) and 8.13(f). The slight difference may be due to the frequency response of each wavelet. The orthogonal wavelet may have wavelet coefficients affected by noise being more highly confined to frequency bands separate from the frequency bands containing wavelet coefficients due mainly to image objects compared with the linear phase wavelet. This is likely, given the scalar power responses shown in Section 6.3.2. However, the effect of the dramatic difference between the scalar power results shown in Section 6.3.2 seems to have been greatly reduced by applying a noise reduction scheme which uses adaptive thresholding across the scales.

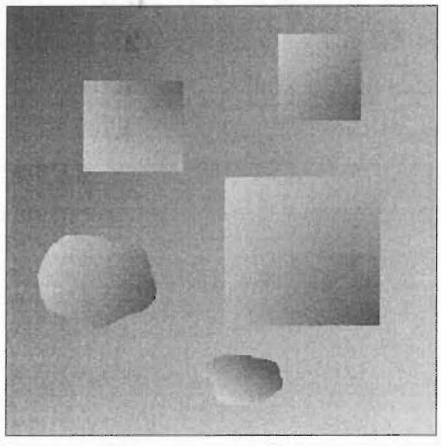
Non-separable wavelets did provide more flexibility for adjustment of enhancement parameters than separable wavelets. There was little difference in the results of the enhancement between the two non-separable wavelets, although the kovaorth wavelet produced slightly better denoising results.



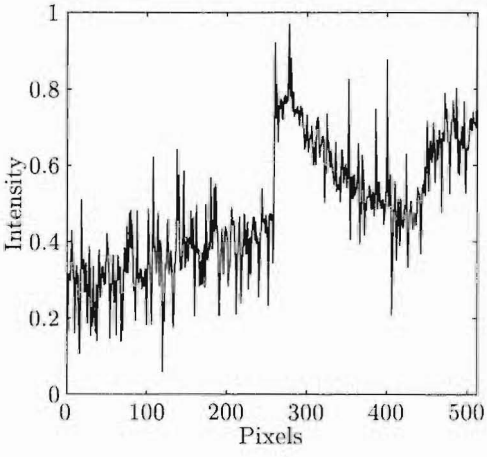
(a) Noisy graduated gray scale image with section of profile.



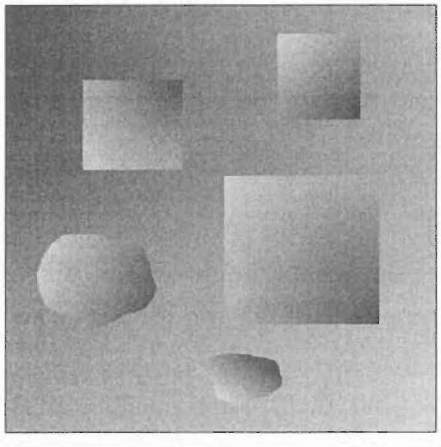
(b) Profile of noisy graduated gray scale image.



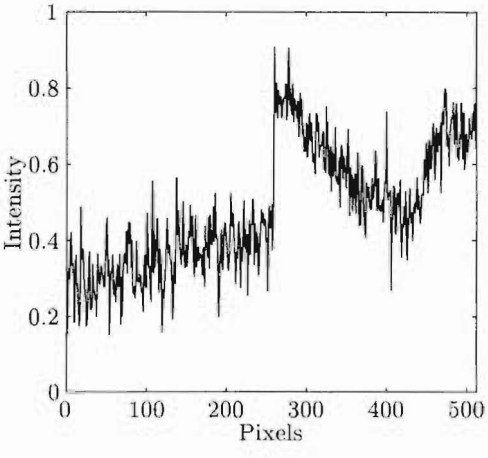
(c) Reconstructed image using non-separable linear phase wavelet.



(d) Profile of (c).



(e) Reconstructed image using non-separable orthogonal wavelet.



(f) Profile of (e).

Figure 8.13 Comparison of image enhancement between non-separable wavelets. (a) Noisy image showing section for profile. (b) Profile of original. (c) Using linear phase wavelet. (d) Profile of (c). (e) Using orthogonal wavelet. (f) Profile of (e).

8.3.2 Comparison Between Enhancement Functions

In this section, three enhancement functions are tested and compared. The enhancement method designed by Laine *et al.* [1995] for enhancing mammography images using the original function and the two enhancement functions proposed in Section 8.2.2 are tested. One of the ultrasound images (300x300 pixels) was used and a range of parameters tested for each of the enhancement methods. The enhancements were performed using the kovaorth wavelet.

The images were assessed subjectively for quality using guidelines from radiologists at Christchurch Hospital. For ultrasound images, several aspects were important, in particular contrast and artefacts. Contrast is related to the the brightness of the objects of interest compared with the brightness of the background. There are two categories of contrast: *gross contrast* refers to the contrast of the main features in an image, such as, in this series of images, the uterine cavity and fallopian tube, compared with the background brightness; *subtle contrast* refers to the contrast of subtle features, such as the uterus wall, compared with the background brightness. It is desirable to have a high intensity for the objects of interest relative to the background. There are also two categories of artefacts: *additional artefacts* and *existing artefacts.* Additional artefacts are features of the enhanced image which are due to the enhancement process and are not related to the object being imaged, and should not be introduced into the image. Finally, for existing artefacts, reduction of the effects of speckle and of beam widening outside the focal zone was considered to be important. Maintaining the effect of the shadow was seen as desirable as it acts as a pointer to objects of interest, such as the uterine cavity.

The method developed by Laine *et al.* [1995] for enhancing images is described in Section 8.2.1. The gain, K, was varied between 1 and 20, with the threshold set as defined. The levels of noise reduction were varied between 1 and 3. Only 256x256 of the image was tested for this method. The results were mixed, with contrast of some of the features being improved, but this was usually accompanied by a worsening of the artefacts. The best result is shown in Figure 8.14 and was obtained using $K = 2$ and noise reduction to 2 levels. In this example, the gross contrast between both the uterine cavity and the fallopian tube and the background in the image is improved. The subtle contrast of the wall of the uterus is easier to see compared with the original. However, blurring of the image features has occurred and the effects of the speckle and beam widening are worse.

The range of the parameters tested for the two new enhancement methods is shown in Table 8.3. The criteria for the enhancement given in Equation (8.9) that K_1 be less than K_2 was relaxed to test the theory. The gain controlling the spatially varying aspect of the enhancement, k, was set to 1 to test the contrast enhancement without spatial adaption. The weight for the denoising threshold, W_d, was applied to the variance

of the decomposition at that scale to give the threshold. The enhancement threshold uses the enhancement weight, W_e, in the equation $T = \bar{c} + W_e\sigma^2$, where the mean, $\bar{c}$, and variance, σ^2, are of the wavelet coefficients at that scale, c. The number of scales denoised, N_d, and enhanced, N_e were also varied. Approximately 12,000 combinations of these parameters were tested.

Parameter	Range of values tested	Range for best images
K_1	0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 0.9, 1, 1.2, 1.5	$0.5 < K_1 < 1.2$
K_2	0.5, 1, 1.5, 2, 2.5, 5	$1 < K_2 < 1.5$, $K_1 < K_2$
k	1, 1.01, 1.02, 1.03, 1.04, 1.05, 1.06, 1.07, 1.08	$k > 1.06$
W_d	0.01, 0.1, 0.2, 0.4, 0.5, 0.6, 0.8, 1, 1.5, 10	
W_e	1–3	W_e = 1–3(d), $W_e = 3$ (a)
N_d	1–3	$N_d = 1$
N_e	1–6	$N_e \leq 4$

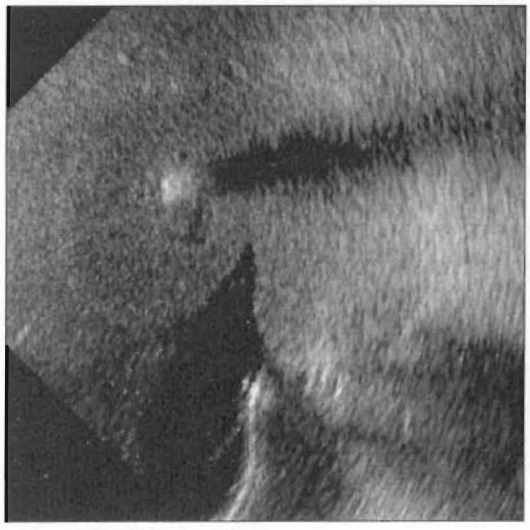
Table 8.3 Range of parameters tested for enhancement techniques, where K_1, K_2, and k are the gains used, W_d is the weight used for the denoising threshold, W_e is the weight for the enhancement threshold, N_d is the number of scales denoised, N_e is the number of scales enhanced.

Contrast enhancement without spatial adaptation was successful, but did not reduce the effects of beam widening, as expected. An example of one of these images is shown in Figure 8.14(c). This image was obtained using parameters: $K_1 = 0.9$, $K_2 = 1.5$, $k = 1$, $W_d = 0.1$, $W_e = 3$, $N_d = 1$, and $N_e = 1$. The gross contrast has been retained with the objects of interest, the uterine cavity and fallopian tube, having good contrast relative to the background. The subtle contrast is slightly improved. No additional artefacts are apparent, however, the effects of the beam widening have only been reduced slightly.

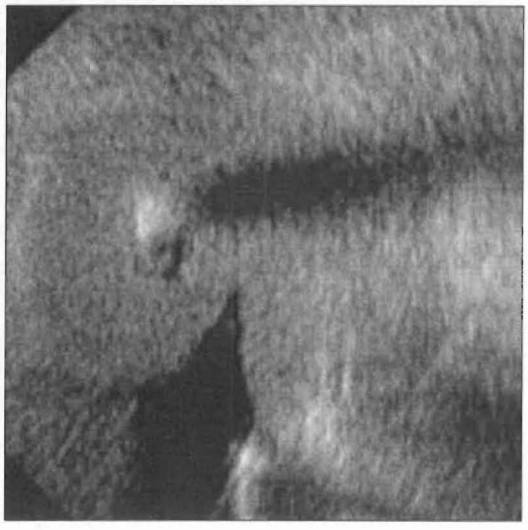
Spatially varying contrast enhancement was found to be the most successful method. The best images had the ranges for the parameters shown in Table 8.3 in the right hand column. An example of one of these images is shown in Figure 8.14(d). This image was obtained using the same parameters as the contrast enhancement image in Figure 8.14(c), but with $k = 1.08$. The gross contrast has been retained with the objects of interest, the uterine cavity and fallopian tube, having good contrast relative to the background. The subtle contrast is slightly improved. No additional artefacts are apparent and the effects of the beam widening have been reduced dramatically.

Of the three image enhancement techniques, the spatially varying contrast enhancement performed the best, followed by the novel contrast enhancement. Both of these performed better than the enhancement developed specifically for the enhancement of mammography images by Laine *et al.* [1995].

In ultrasound images, beam widening outside the focal zone is the most visually disturbing artefact. The new technique, spatially varying contrast enhancement, was the only method to effectively deal with this problem by reducing the intensity of the pixels affected by beam widening. The reduction in intensity is performed in proportion

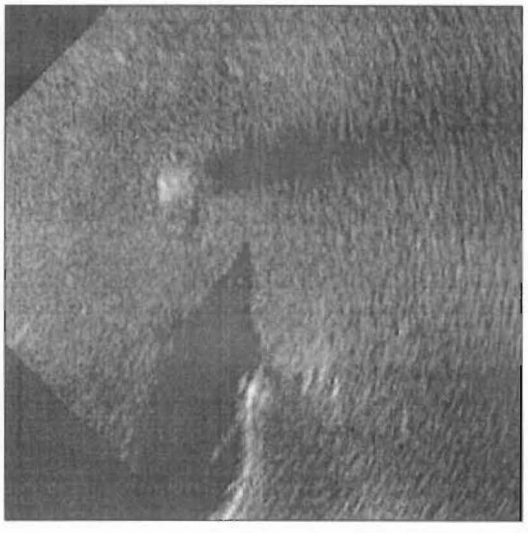


(a) Original ultrasound image



(b) Enhancement of the ultrasound image using the technique developed by Laine.

(c) Enhanced ultrasound image using novel contrast enhancement technique.



(d) Enhanced ultrasound image using novel spatially varying technique.

Figure 8.14 Results of enhancement using various techniques. (a)Original ultrasound image (300x300). (b) Image enhanced using technique from Laine *et al.* [1995]. This image is 256x256 cropped from the original. (c) Image enhanced using contrast enhancement obtained using parameters: $K_1 = 0.9$, $K_2 = 1.5$, $k = 1$, $W_d = 0.1$, $W_e = 3$, $N_d = 1$, and $N_e = 1$. (d) Image enhanced using spatially varying contrast enhancement with the same parameters as (c), except for $k = 1.08$.

to the distance from the focal zone and, therefore, to the amount that the pixels were widened. Since this method was successful, it was applied to a range of ultrasound images to test the performance further.

8.3.3 Consistency of Spatially Varying Contrast Enhancement Function

Spatially varying contrast enhancement was applied to all the ultrasound images in the volume obtained using the phantom from Christchurch Hospital (see Section 1.3). The volume is comprised of 54 consecutive two-dimensional slices, each slice being one image. For the first test, the non-separable orthogonal wavelet was applied to all the images, with the same parameters as used for the image in Figure 8.14(d). The parameters were $K_1 = 0.9$, $K_2 = 1.5$, $k = 1.08$, $W_d = 0.1$, $W_e = 3$, $N_d = 1$, and $N_e = 1$.

All the images showed similar improvements. The gross and subtle contrasts were maintained or improved. The artefact due to beam widening was reduced and no additional artefacts were introduced. Two of the images are shown in Figure 8.15 with the original on the left and the enhanced images on the right.

Two further tests of five different parameter settings for each test were done for separate subsets of the images. Each image subset was comprised of six images that were taken at variable spacings through the volume of images. The parameter settings were chosen to produce variable image quality results. Again, the images showed similar results for given parameter settings.

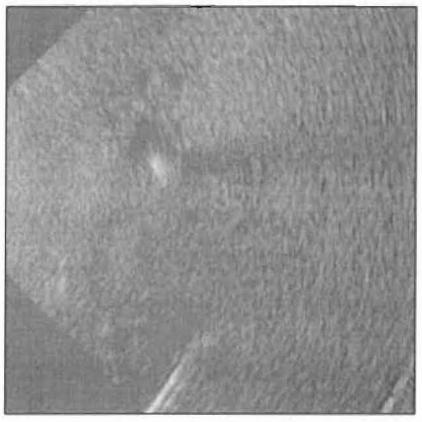
8.3.4 Discussion of Image Enhancement Functions

Comparison of the enhancement functions was performed using subjective methods. The criteria for assessing the different images were developed in discussion with experts. However, the problems discussed in Chapter 7 apply here. That is, the quality of images is difficult to assess when different characteristics are conflicting in quality, also, repeatability of assessment is not guaranteed between observers, or even for one observer at different times. However, the subjective assessments were tested using a number of observers, and the assessments of images that were significantly better than the original were consistent. The development of an objective measure of image quality would greatly improve consistency and, if fully automated, could be used in conjunction with the methods presented here to iteratively improve image quality.

Four categories of wavelets were compared for enhancing test images. The non-separable wavelets allowed a greater customisation of the enhancement process, as expected [Kovačević and Vetterli 1995]. A greater difference might have been expected between the non-separable wavelets because of the difference shown in the power at each scale in Section 6.3.2, however, using the adaptive thresholds seems to reduce the differences between the wavelets.

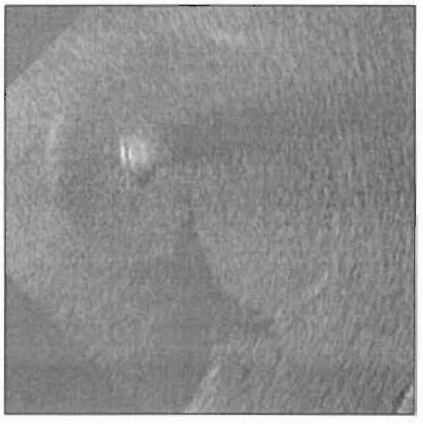
The functions presented here for contrast enhancement and spatially varying contrast enhancement performed better than the function designed for enhancing mammography images [Laine *et al.* 1995]. It is thought that the novel functions performed

(a) Original image of slice 16



(b) Enhanced image of slice 16

(c) Original image of slice 31



(d) Enhanced image of slice 31

Figure 8.15 Two original images (left), with their counterparts enhanced using spatially varying contrast enhancement.

better as they were adapted to ultrasound images. Mammography images are of high intensity and the features of interest are close in intensity to the background. In comparison, ultrasound images generally have a low intensity background in the focal zone, with relatively high contrast features. The enhancement functions developed in this thesis were designed to take advantage of that difference, ensuring the wavelet coefficients due to the edges of objects, which have high values, were enhanced. The wavelet coefficients due to noise were generally of smaller magnitude and were therefore further depressed using the new techniques.

The function used for spatially varying the contrast enhancement was able to effectively reduce the effects of beam widening because the effect of the artefact increases linearly with distance. Spatially varying contrast enhancement can be applied to dif-

ferent images from the same volume using set parameters with success. The technique was not tested on images obtained using different scanner settings or from different scanners. It is expected that the parameters would need to be modified for different scanner settings or machines, but nevertheless, the principle of spatially varying the contrast enhancement should be portable.

8.4 CONCLUSIONS

Separable transforms decompose an image into three detail components at each scale, each exhibiting a directional bias in either the horizontal, vertical, or diagonal direction. Non-separable wavelet transforms produce a single detail component that contains all the directional information at each scale. The final decomposition level in both cases contains an approximation component that represents low frequency information of the images. If the directional biases of the separable transform are to be avoided then a non-separable wavelet transform should be used.

Several methods of image enhancement are described using data manipulation in the wavelet domain. Methods of noise reduction that were computationally efficient were considered. Noise reduction using hard- and soft-thresholding is described. Two new methods of denoising are introduced: Wavelet Median Denoising and Two-Dimensional Derivative Denoising, an extension of a one-dimensional scheme. Three methods of noise reduction, Two-Dimensional Derivative Denoising, soft thresholding denoising and Wavelet Median Denoising, were compared. The first method, which was a novel extension of a one-dimensional method [Roy *et al.* 1999], did not meet the preliminary conditions for success, so was not tested further. The second method, soft thresholding [Donoho 1995], and the third method, Wavelet Median Denoising, were applied to two test images for a range of values for each parameter. Noise reduction applied to increasing scales of the wavelet transform produced successively more blurred reconstructed images. At lower scales, where the blurring was not significant, Wavelet Median Denoising performed as well as or better than soft thresholding for the non-separable wavelets and the separable orthogonal wavelet.

Three different enhancement functions are presented. The first function, from Laine *et al.* [1995], increases low contrast areas more than high contrast areas. The second function, Contrast Enhancement, was developed for application to ultrasound images and increases the contrast where objects are detected and reduces the contrast in the background. The third function, Spatially Varying Contrast Enhancement, adapts the second function in order to spatially vary contrast in the image. The enhancement function is spatially varied because one of the artefacts in ultrasound images is spatially dependant. The last two enhancement functions are original and are designed specifically for the enhancement of ultrasound images. All three contrast enhancement functions are assessed for their capacity to increase the quality of ultrasound images.

The original functions developed in this thesis produced better contrast enhancement and reduction of unwanted features in the image compared with the function from Laine *et al.* [1995], with the spatially varying contrast enhancement function performing best. Little difference was found between the performances of the wavelets, probably due to the adaptive nature of the thresholds across scales and their similar scalar power responses. The two non-separable wavelets did, however, provide more flexibility in designing the enhancement functions. Therefore, to enhance ultrasound images of the type depicted here, of the presented techniques, the best technique is to use Laine's technique for enhancement with Wavelet Median Denoising and the spatially varying contrast enhancement function, calculated using the kovaorth wavelet.

A given set of parameters for spatially varying contrast enhancement produce similar results for images taken at the same settings on the same machine. This suggests that parameters can be found which enhance the images found for one scanner, greatly simplifying the application of the algorithm.

Chapter 9

CONCLUSIONS

In this thesis, one-dimensional wavelets have been applied to infant breathing signals to quantify amplitude modulation, and two-dimensional wavelets have been applied to enhance ultrasound images. A new technique, the *Wavelet Performance Measure*, was used to select a wavelet to be applied to infant breathing signals. Two novel measures, *rhythmicity* and *impact*, were defined to assess the extent of amplitude modulation in the breathing signals and the effect of the amplitude modulation on the primary breathing signal, respectively. Using rhythmicity and impact, differences were found between infants at high risk for the Sudden Infant Death Syndrome (SIDS) and infants at low risk for SIDS, and between infants who later succumbed to SIDS and their controls.

An *Image Quality Equation* was developed for assessing the quality of ultrasound images. Two novel methods for noise reduction, *Two-Dimensional Derivative Denoising* and *Wavelet Median Denoising*, were presented. The quality of ultrasound images was also improved using two new functions for image enhancement, *Contrast Enhancement* and *Spatially Varying Contrast Enhancement.*

This chapter presents conclusions drawn from the development and application of these methods and presents directions for further research.

9.1 APPLICATION OF ONE-DIMENSIONAL WAVELETS

Wavelets were applied to the analysis of infant breathing signals to measure amplitude modulation, a breathing pattern that has had little investigation, partly due to the lack of appropriate measuring techniques. Infant breathing signals are widely studied, particularly when considering Sudden Infant Death Syndrome (SIDS). It has been postulated that infants at high risk for SIDS have physiological mechanisms that work differently from infants at low risk for SIDS. Studies of breathing signals show patterns and characteristics which, when studied in conjunction with other systems in the body, may lead to an understanding of the physiological mechanisms surrounding SIDS. This understanding may lead to possible intervention that could be useful in preventing

cases of SIDS. Amplitude modulation is related to periodic breathing, a pattern that has been associated with SIDS and infants at high risk for SIDS, and hence is worthy of further investigation.

The selection of a wavelet to analyse infant breathing signals was achieved using a new measure, the Wavelet Performance Measure. The measure was based on the performance of wavelets applied to test signals in the wavelet domain, assessed using the energy in the wavelet domain. Wavelets were ranked higher if they were localised in the wavelet domain in both temporal and scalar dimensions and if they had rapidly decaying scalar harmonics. The ranking of wavelets using the Wavelet Performance Measure is different from other traditional methods of ranking wavelets [Daubechies 1992, Unser and Aldroubi 1996, Mojsilovic *et al.* 1998]. Using the Wavelet Performance Measure, the symlet wavelet of order 8 (sym8) was shown to be the most appropriate wavelet for analysing infant breathing, of the wavelets that were assessed. Traditional methods for selecting wavelets, such as the number of vanishing moments and the lower bound for the Heisenburg Uncertainty Principle, gave different optimum wavelets. However, when wavelet transforms for the different wavelets were compared, this confirmed sym8 gave the best transform, and hence the Wavelet Performance Measure selected the most appropriate wavelet for this application. Unlike traditional methods that only measure characteristics of the wavelets themselves, the wavelet is not selected in isolation but in conjunction with the signal to be analysed and the salient properties to be estimated. A second reason for the improved performance of the Wavelet Performance Measure is that test signals are used that are specific to the application.

Having selected a wavelet, breathing signals were analysed in the wavelet domain. Infant breathing was shown to contain non-stationary frequency components, and therefore, as expected, analyses using wavelets were shown to be more appropriate than Fourier based techniques for analysing the patterns of interest. Wavelets are not necessarily the most appropriate tool for analysing all breathing patterns. For example, breathing rate and breathing rate variability have been successfully measured using peak detection [Schechtman *et al.* 1988]. Using wavelet analysis, frequency components can be isolated in both temporal and scalar dimensions, and hence this technique is suited to detecting patterns characterised by a frequency component, occurring during distinct periods of breathing—such as amplitude modulation.

Rhythmicity, a novel measure for assessing amplitude modulation in infant breathing signals during quiet sleep, was introduced. Previously, the only studies of this pattern of breathing used Fourier analysis [Roy *et al.* 1999]. Rhythmicity was shown to successfully assess the extent of constant amplitude frequency components with significant magnitude in the breathing signals. A second measure, impact, was shown to successfully assess the effect of amplitude modulation on the principal breathing frequency. Using rhythmicity for assessing breathing that showed significant amounts of amplitude modulation, infants at high risk for SIDS were shown to have greater

constant frequency amplitude modulation than infants at low risk for SIDS. Infants that later succumbed to SIDS were shown to have a greater impact of amplitude modulation in breathing that showed significant amounts of amplitude modulation than their controls. For breathing that was as regular as possible, both infants who later succumbed to SIDS and their controls exhibited significant amounts of rhythmicity of amplitude modulation, however, there was a significant time of night effect. Control infants tended to exhibit less amplitude modulation later in the night, while infants who later succumbed to SIDS tended to exhibit more amplitude modulation later in the night. This finding suggests that SIDS and control infants have differences in their respiratory control and that wavelets are useful in illustrating that difference.

9.2 APPLICATION OF TWO-DIMENSIONAL WAVELETS

B-scan ultrasound scanners are widely used to produce images of soft tissues in the human body. They are non-invasive and the process is not harmful to human tissue, providing safety guidelines are followed. The images, however, are prone to artefacts, most notably speckle and beam widening. Often, the features of interest have an intensity that is close to the background intensity, making them hard to distinguish. In this thesis, the aim was to reduce artefacts while increasing the contrast between the features of interest and the background of the image. This aim was achieved using two-dimensional wavelet transforms in conjunction with enhancement techniques in the wavelet domain. These techniques were designed specifically for ultrasound images, and did improve image quality, confirming the usefulness of wavelets for ultrasound image enhancement.

Wavelets in two dimensions may be constructed either separably, using the tensor product of one-dimensional wavelets, or non-separably. Separable wavelet transforms may be performed separately in each dimension, whereas, non-separable wavelet transforms are performed simultaneously in all dimensions. The transform of the separable wavelet necessarily results in four separate images for each level of transform in the wavelet domain. Each level is downsampled by four from the previous level. Non-separable wavelets, however, result in a more holistic transform, producing only one image per level in the wavelet domain, making the transform easier to interpret. The use of the quincunx sub-sampling lattice results in downsampling by two, giving finer frequency bandwidths for each level. Narrower bandwidths for each level allow a greater freedom in designing image enhancement techniques, as different features are spread into separate levels meaning that those features can be treated separately in the enhancement.

A problem in assessing image quality is finding a reliable, objective measure. For this reason, an image quality equation was developed to objectively measure the image quality of images. However, despite the parameters for the equation each measuring

a potential aspect of image quality, the equation did not predict image quality consistently. Investigations revealed several contributing factors. Intra- and inter-personal discrepancies in rating test images, possibly resulting from lack of sufficient training, may have had some bearing on the result. Given that the gold standard of image quality is most often human judgment, inconsistent rankings by observers means a quality measure cannot provide an exact fit. It is also possible that the parameters used for the equation were not appropriate. For instance, the power of the intensity of the image was expected to drop with increasing image quality, due to a reduction in noise, however, this was not the case. It is possible that the shape of the one-dimensional approximation to the power spectra is a more accurate indicator of image quality. Peaks in the spectra may indicate that certain frequency components have been overly increased in the enhancement procedure, relative to other frequency components. It can be concluded that each image quality parameter must be tested to ensure that it is indeed a predictor of image quality.

Two techniques for reducing the noise in images were presented. The first of the novel denoising techniques presented, Two-Dimensional Derivative Denoising, did not meet the primary condition for success that peaks in the scalar power due to noise and peaks due to image features should be separated further apart for the derivative than the original. It is possible that the technique would be successful for a larger image, or an image with higher resolution, and different spectral characteristics than the images used here. It is also possible that different wavelets or improved estimates for the derivative and integration could produce a successful result. The second new denoising technique presented, Wavelet Median Denoising, had comparable results to soft threshold denoising [Donoho 1995] applied to low scales. In some cases, Wavelet Median Denoising showed a larger improvement in the signal to noise ratio than soft threshold denoising. Both Wavelet Median Denoising and soft threshold denoising introduced blurring at higher levels of denoising. Wavelet Median Denoising is computationally more expensive than soft thresholding.

Of the methods tested for enhancement of ultrasound images, the best method was the novel spatially varying contrast enhancement method using the kovaorth wavelet and Wavelet Median Denoising. For a given set of parameters, the method was shown to enhance a number of images recorded using the same scanner and scanner settings with similar results. The main artefacts in ultrasound images are beam widening and speckle noise, which were reduced using this technique, therefore, this technique may be successful for other settings and scanners.

9.3 SUGGESTIONS FOR FURTHER RESEARCH

In this section, suggestions for future areas of investigation deriving from research in this thesis are proposed.

In the evaluation of rhythmicity, an averaging operation takes place when calculating scalar energy which could result in some transient frequency components being missed. It would be possible to improve the scalar energy calculation using an iterative approach, however, this would be at the expense of computational efficiency. Improving the calculation of scalar energy could result in shorter or smaller magnitude frequency components being included in the assessment, leading to a more accurate result in terms of time–frequency localisation.

While the results in this thesis suggest that amplitude modulation shows differences between SIDS and non–SIDS groups, this pattern could be further investigated from a physiological perspective. A larger study, including more infants and epochs for each infant, could be analysed. Studies examining the interaction between the physiological systems could also be carried out. The timing of the occurrence of frequency components then becomes more important and, therefore, even more appropriate for non-stationary signal analyses, such as those used in this thesis.

Although only amplitude modulation of infant breathing signals has been considered, other physiological signals also exhibit low frequency components. Cardiac signals, which are generally of a higher frequency than respiratory signals, often have modulation effects due to respiration and other low frequency components known as Mayer waves [Preiss *et al.* 1975]. Rhythmicity could be applied to these signals to quantify the effects of the low frequency components.

If the image quality equation is improved to consistently rank the quality of images, and the assessment of all the parameters is automated, then it may be possible to construct an iterative approach to image enhancement using the enhancement methods presented. Two steps are suggested: firstly the development of a gold standard of images ranked by radiologists, as they are ultimately the end–users of any ultrasound image enhancement scheme. Secondly, the determination of parameters that are predictive of image quality as defined by the gold standard.

The enhancement techniques were tested on one volume of images obtained from a phantom. Whilst this was adequate for developing the algorithm, further studies on different ultrasound images would test the applicability of the method for a wider range of images. Ideally, images from different scanners and of live subjects would be used.

The noise reduction technique, Wavelet Median Denoising, could probably be improved with the implementation of an adaptive criterion for denoising. The local properties of an image could determine the neighbourhood used for the median filtering, or the percentile value could be changed depending on the local characteristics. For example, if the neighbourhood was over an edge, then a high percentile value could be used, but if the neighbourhood was over a nearly homogeneous background, then a lower percentile value could be used.

Another suggestion for further research is the three-dimensional visualisation of

two-dimensional wavelet transforms, allowing the visualisation of continuous wavelet transforms of images. It is possible that features in the two-dimensional wavelet transform will be easier to interpret in the continuous wavelet transform than in the discrete wavelet transform, as it is in the one-dimensional case.

Radiographers prefer to be able to analyse ultrasound images in real time. Developing fast algorithms using the enhancement techniques developed in this thesis would provide a valuable tool for radiographers interpreting ultrasound images.

The application of three-dimensional non-separable wavelet transforms to volume data is also possible. An extension of the quincunx sub-sampling scheme may be used for non-separable sub-sampling by two in three dimensions. The extension is known as *face centred orthorhombic* (fco) sub-sampling and has been used in coding schemes for video [Vetterli *et al.* 1990, Kovačević and Vetterli 1993]. Kovačević and Vetterli [1995] also develop a three-dimensional non-separable wavelet using the cascade presented in Section 6.2.2.2. The volume of ultrasound slices examined in this thesis could be analysed using the three-dimensional wavelet and fco sub-sampling. The enhancement techniques developed here could be extended to three dimensions and it is possible that better results than those obtained, in this thesis could be obtained given the extra degree of freedom in three dimensions.

Another area of research is the three-dimensional visualisation of volumes of two-dimensional enhanced images. One of the main problems in visualising ultrasound in three dimensions is the amount of noise in the image. Using the techniques described and developed in this thesis, a series of images could be enhanced, producing much clearer images that could then be used to construct a three-dimensional image. Cleaner images would allow better rendering of surfaces in three dimensions.

Summarising, this thesis presents the development, selection and application of wavelets for the analysis of infant breathing signals and the analysis of ultrasound images. A novel method to select a wavelet for the analysis of infant breathing is introduced, as well as two new measures, based on the energy in the wavelet domain, for assessing amplitude modulation of breathing. A study of two groups of infants, one group at low risk for Sudden Infant Death Syndrome (SIDS), and the second group at high risk for SIDS, showed a predominance of constant amplitude modulation in the group at high risk for SIDS. Another study of infants who later succumbed to SIDS and their controls demonstrated a predominance of constant amplitude modulation for infants who later succumbed to SIDS of breathing showing large amounts of amplitude modulation. For breathing showing little, or no amplitude modulation, a significant time of night effect was noted with control infants tending to decrease rhythmicity due to amplitude modulation during the night, while infants who later succumbed to SIDS showed an increase. For the analysis of ultrasound images, an image quality equation

was developed. However, the results were inconsistent, therefore, subjective criteria for the assessment of image quality were used. Noise reduction and contrast enhancement techniques were developed that are used in the wavelet domain to improve the quality of ultrasound images. Using the performance criteria, non-separable wavelets in conjunction with the enhancement techniques developed in this thesis produced reconstructed images that had higher contrast and significantly less effects from existing artefacts than the original images.

In conclusion, wavelets are a useful tool in the analysis of one-dimensional and two-dimensional medical signals. Wavelets can improve the quality of information available from medical signals and could become part of standard techniques used for analysing those signals.

Appendix A

CUMULATIVE METHOD FOR INTEGRATING DERIVATIVES

This section describes and defines a method of integrating derivatives. The proposed cumulative method is based on a linear approach.

Given a one-dimensional curve, $f(x)$, $x \epsilon \mathcal{R}$, then the curve may be approximated using a linear piecewise curve, $f[n]$, where $n \epsilon \mathcal{Z}$. The derivative between each point on $f(x)$ is approximated by the slope of $f[n]$ between the points (see Figure A.1).

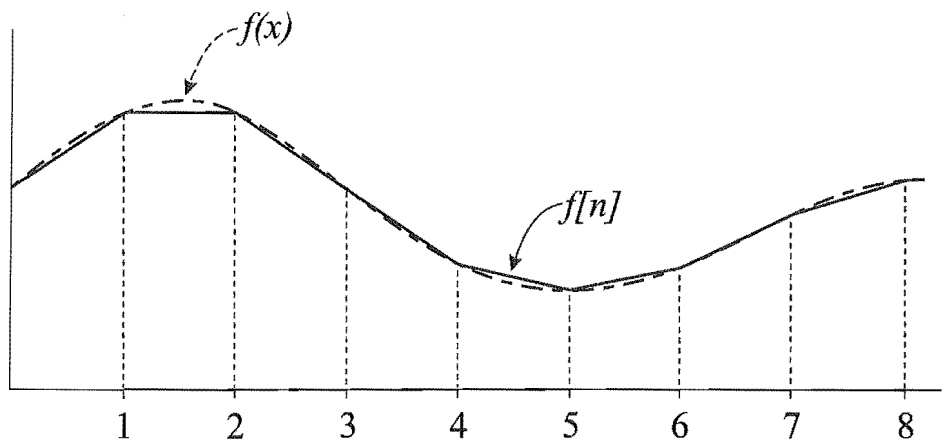


Figure A.1 A curve $f(x)$ (dashed line) is approximated by a linear piecewise curve $f[n]$ (solid line).

The slope $f'[n]$ of $f[n]$ between any two consecutive points can be estimated by the forward finite difference:

$$f'[i+1] = \frac{f[i+1] - f[i]}{(i+1) - (i)} = f[i+1] - f[i]. \tag{A.1}$$

It is then possible to obtain $f[n]$ from $f'[n]$ if the first point $f[1]$ is known. Simply:

$$f[i+1] = f[i] + f'[i+1] = f[i] + f[i+1] - f[i]. \tag{A.2}$$

This concept is expanded into two dimensions. Let X be a surface which is sampled at integer spacing in orthogonal dimensions, x and y. This can be represented by a matrix:

$$X[i,j] = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix} \tag{A.3}$$

where $i, j \in \mathcal{Z}$ are the indices of the sampling points and $a, b, c, \ldots$ are the values of the surface at the sampling points.

Now, the derivative in each dimension is estimated:

$$\frac{dX}{dx} \approx X'_i[i+1,j] = X[i+1,j] - X[i,j] \text{ and } \frac{dX}{dy} \approx X'_j[i,j+1] = X[i,j+1] - X[i,j]. \tag{A.4}$$

In the example, this becomes:

$$X'_i[i,j] = \begin{pmatrix} 0 & b-a & c-b \\ 0 & e-d & f-e \\ 0 & h-g & k-h \end{pmatrix} \tag{A.5}$$

$$X'_j[i,j] = \begin{pmatrix} 0 & 0 & 0 \\ d-a & e-b & f-c \\ g-d & h-e & k-f \end{pmatrix}, \tag{A.6}$$

with zero padding used to retain the original dimensions of the array.

To obtain the original matrix, the first row and column is straightforward and can be conducted as for the one-dimensional case with $X[1,1]$ known, $\hat{X}[1,j+1] = X[1,j] + X'_i[1,j+1]$ and $\hat{X}[i+1,1] = X[i,1] + X'_j[i+1,1]$. Now there is a choice as to how to proceed since each element in the matrix may be derived from two separate element and slope combinations, that is $\hat{X}_i[i+1,j+1] = X[i+1,j] + X'_i[i+1,j+1]$ or $\hat{X}_j[i+1,j+1] = X[i,j+1] + X'_j[i+1,j+1]$. In the ideal case, $\hat{X}[i+1,j+1]$ will be the same for both choices. However, since the two choices for the application of integration after Derivative Denoising may not give the same results and there is no reason to take one over the other, the average of the two possible solutions is taken.

For the example, this leads to the solution:

$$\hat{X}[i,j] = \begin{pmatrix} a & a+(b-a) & a+(b-a)+(c-b) \\ a+(d-a) & \hat{X}[2,2] & \hat{X}[2,3] \\ a+(d-a)+(g-d) & \hat{X}[3,2] & \hat{X}[3,3] \end{pmatrix} \tag{A.7}$$

where

$$\begin{aligned}
\hat{X}[2,2] &= \frac{(a+(d-a)+(e-d)) \ + \ (a+(b-a)+(e-b))}{2}, \\
\hat{X}[2,3] &= \frac{(\hat{X}[2,2] \ + \ (f-e))+(a+(b-a)+(c-b)+(f-c))}{2}, \\
\hat{X}[3,2] &= \frac{(a+(d-a)+(g-d)+(h-g)) \ + \ \hat{X}[2,2]}{2}, \\
\hat{X}[3,3] &= \frac{(\hat{X}[3,2]+(k-h))+(\hat{X}[2,3]+(k-f))}{2},
\end{aligned}$$

and in the ideal case, $\hat{X} = X$.

REFERENCES

Akay, M. and Mulder, E. (1996), "Investigating the Effect of Maternal Alcohol Intake on Human Fetal Breathing Rate Using Adaptive Time-Frequency Analysis Methods.," *Early Human Development*, Vol. 46, No. 1-2, September 1996, pp. 153–164.

Akay, M. and Szeto, H.H. (1995), "Investigating the Relationship Between Fetus EEG, Respiratory and Blood Pressure Signals During Maturation Using Fast Wavelet Transform," *Annals of Biomedical Engineering*, Vol. 23, No. 5, September/October 1995, pp. 574–582.

Akay, M., Akay, Y. and Szeto, H. (1996), "The Effects of Morphine on the Relationship Between Fetal EEG, Breathing and Blood Pressure Signals Using Fast Wavelet Transform," *Biological Cybernetics*, Vol. 74, No. 4, April 1996, pp. 367–372.

Anant, K. and Dowla, F. (1995b), "Wavelet transform methods for phase identification in three-component seismograms," *The Bulletin of the Seismological Society of America*, Vol. 87, No. 6, December 1995b, pp. 1598–1612.

Anant, K., Dowla, F. and Rodrigue, G. (1994), "Detection of the electrocardiogram P-wave using wavelet analysis," *Proceedings of the SPIE*, Vol. 2242, April 1994, pp. 744–749.

Anant, K., Dowla, F. and Rodrigue, G. (1995a), "Vector quantisation of ECG wavelet coefficients," *IEEE Signal Processing Letters*, Vol. 2, No. 7, July 1995a, pp. 129–131.

Anderson, N. (1994), Private correspondence, 1994.

Anton, H. (1984), *Elementary Linear Algebra*, John Wiley and Sons, Drexel University, 4^{th} ed., 1984.

Ariagno, R., Guilleminault, C., Korobkin, R., Owen-Boeddiker, M. and Baldwin, R. (1983), "'Near-miss' for sudden infant death syndrome infants: a clinical problem," *Pediatrics*, Vol. 71, No. 5, 1983, pp. 726–730.

Aussem, A., Campbell, J. and Murtagh, F. (1998), "Wavelet-based feature extraction and decompostition strategies for financial forecasting," *Journal of Computational Intelligence in Finance*, Vol. 6, 1998, pp. 5–12.

Bentley, P. and McDonnell, J. (1994), "Wavelet Transforms: an Introduction," *Electronics and Communication Engineering Journal*, August 1994, pp. 175–186.

Bloomfield, P. (1976), *Fourier Analysis of Time Series: An Introduction*, John Wiley & Sons, Inc., USA, 1976.

Böröczky, L., Fioravanti, R., Fioravanti, S. and Giusto, D. (1995), "Speckle Noise Filtering in SAR Images Using Wavelets," In Carlo Braccini, L.D. and Veernazza, G. (Eds.),*Image Analysis and Processing*, San Remo, Italy, September 13-15 1995, pp. 671–676.

Bracewell, R.N. (1986), *The Fourier Transform and its Applications*, McGraw-Hill, New York, 2nd rev. ed., 1986.

Bradley, J.N. and Brislawn, C.M. (1994a), "The Wavelet/Scalar Quantisation Compression Standard for Digital Fingerprint Images," In *Proceedings of IEEE ISCAS-94*, May 1994a, pp. 205–208.

Bradley, J.N., Brislawn, C.M. and Hopper, T. (1994b), "The FBI Wavelet/Scalar Quantization Fingerprint Image Compression Standard," In *Proceedings of the National Media Laboratory Conference Solid-State Memory*, Pasadena, California, USA, 1994b, pp. A11–A14.

Brooks, J. (1992), "Apparent life-threatening events and apnea of infancy," *Clinics in Perinatology*, Vol. 19, No. 4, 1992, pp. 809–838.

Brown, P., Dove, R., Tufnell, C. and Ford, R. (1992), "Oscillations of Body Temperature at Night," *Archives of Disease in Childhood*, Vol. 67, 1992, pp. 1255–1258.

Cen, S., Persson, H., Schilling, D., Cosman, P. and Berry, C. (1998), "Human Observer Responses to Progressively Compressed Images," In *Conference Record of the Thirty-First Asilomar Conference on signals, Systems and Computers*, IEEE Computer Society Press, Los Alamitos, CA, USA, 1998, pp. 657–661.

Chan, H.L., Lin, J.L., Du, C.C. and Wu, C.P. (1997), "Time-Frequency Distribution of Heart Rate Variability Below 0.05 Hz by Wigner-Ville Spectral Analysis in Congestive Heart Failure Patients," *Medical Engineering and Physics*, Vol. 19, No. 6, September 1997, pp. 581–587.

Chang, S.G. and Vetterli, M. (1997), "Spatial Adaptive Wavelet Thresholding for Image Denoising," In *Proceedings of the IEEE International Conference on Image Processing*, 1997, pp. 374–377.

Chen-Li, Nguyen-Truong-Q and Chan-Kwok-Ping (1995), "Symmetric Extension Methods for M-Channel Linear-Phase Perfect-Reconstruction Filter Banks," *IEEE Transactions on Signal Processing*, Vol. 43, No. 11, November 1995, pp. 2502–2511.

Comes, S., Bruyndonckx, O. and Macq, B. (1995), "Image Quality Criterion Based on the Cancellation of the Masked Noise," In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing - 95*, IEEE, New York, NY, USA, 1995, pp. 2635–2638.

Cosman, P.C., Tseng, C., Gray, R.M., Olshen, R.A., Moses, L.E., Davidson, H.C., Bergin, C.J. and Riskin, E.A. (1993), "Tree-Structured Vector Quantization of CT Chest Scans: Image Quality and Diagnostic Accuracy," *IEEE Transactions on Medical Imaging*, Vol. 12, No. 4, December 1993, pp. 727–739.

Cosman, P.C., Gray, R.M. and Olshen, R.A. (1994), "Evaluating Quality of Compressed Medical Images: SNR, Subjective Rating, and Diagnostic Accuracy," *Proceedings of the IEEE*, Vol. 82, No. 6, June 1994, pp. 919–932.

Croisier, A., Esteban, D. and Galand, C. (1976), "Perfect channel splitting by use of interpolation, decimation, tree-decomposition techniques," In *Proceedings of the International Conference on Information Science/Systems*, Patras, Greece, August 1976, pp. 443–446.

Crouse, M.S., Nowak, R.D. and Baranuik, R. (1998), "Wavelet-Based Statistical Signal Processing Using Hidden Markov Models," *IEEE Transactions on Signal Processing*, Vol. 46, No. 4, April 1998, pp. 886–902.

Daubechies, I. (1992), *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992.

Daubechies, I. (1993), "Wavelet Transforms and Orthonormal Wavelet Bases," In Daubechies, I. (Ed.),*Different Perspectives on Wavelets*, San Antonio, Texas, January 1993, pp. 1–34.

Dehkordi, M., Erfanian, A. and Foroutan, A. (1999), "Time-Frequency Analysis of the Heart-Rate Variability During Physical and Physical-Pharmacological Tests," In *Proceedings of the First Joint BMES/EMBS Conference*, IEEE Engineering in Medicine and Biology and Biomedical Engineering Society, October 1999, P. 261.

Donoho, D.L. (1995), "De-noising by Soft-Thresholding," *IEEE Transactions on Information Theory*, Vol. 41, No. 3, May 1995, pp. 613–627.

Dove, R., Brown, J., Fright, R., Tuffnell, C. and Ford, R. (1990), "Computer Polygraphic System for Infants at Risk for Sudden Infant Death Syndrome (SIDS),"

Australasian Physical and Engineering Sciences in Medicine, Vol. 13, 1990, pp. 188–191.

Driggers, R.G., Cox, P.G., Leachtenauer, L., Vollmerhausen, R. and Scribner, D.A. (1998), "Targeting and Intelligence Electro-Optical Recognition Modeling: A Juxtaposition of the Probabilities of Discrimination and the General Image Quality Equation," *Optical Engineering*, Vol. 37, No. 3, March 1998, pp. 789–797.

Drouiche, K. and Kateb, D. (1999), "New Filter Banks and More Regular Wavelets," *IEEE Transactions on Signal Processing*, Vol. 47, No. 8, August 1999, pp. 2220–2227.

Dunne, K.P., McKay, M. and Matthews, T.G. (1986), "'Near miss' sudden infant death and obstructive apnoea," *Archives of Disease in Childhood*, Vol. 61, 1986, pp. 1039–1040.

Engeldrum, P.G. (1995), "A Framework for Image Quality Models," *Journal of Imaging Science and Technology*, Vol. 39, No. 4, July - August 1995, pp. 312–318.

Esteban, D. and Galand, C. (1977), "Application of quadrature mirror filters to split-band voice coding schemes," In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Washington, DC., 1977, pp. 191–195.

Ford, R.P.K. (1986), "Postneonatal mortality in Christchurch," *The New Zealand Medical Journal*, Vol. 99, No. 815, 1986, pp. 939–941.

Ford, R., Brown, P., Dove, R., Tufnell, C. and Macey, P. (1992), "Homelog: Long Term Recording of Infant Temperature, Respiratory and Cardiac Signals in the Home Environment," *Journal of Paediatrics and Child Health*, Vol. Suppl. 1, 1992, pp. 26–33.

Ford, R.P.K., Tuffnell, C.S., Macey, P.M., Tappin, T.M. and Sambamoorthy, M. (1996), "Rectal Temperature Changes and Apnea," In *Conference program of the Fourth SIDS International Conference*, SIDS International, Washington, USA, 1996, P. 123.

Gabor, D. (1946), "Theory of Communication," *Journal of IEE*, Vol. 93, 1946, pp. 429–457.

Gibson, E. (1996a), *Intensive Care of the Fetus and Neonate* St. Louis: Mosby-Year Book, 1996a, ch. Apnea, P. 470.

Gibson, E. (1996b), *Intensive Care of the Fetus and Neonate* Mosby-Year Book, Inc, St. Louis, USA, 1996b, ch. Sudden infant death syndrome, pp. 482–493.

Gordon, D., Cohen, R.J., Kelly, D., Akselrod, S. and Shannon, D.C. (1984), "Sudden Infant Death Syndrome: abnormalities in short term fluctuations in heart rate and respiratory activity," *Pediatric Research*, Vol. 18, No. 10, 1984, pp. 921–926.

Gordon, D., Southall, D., Kelly, D., Wilson, A., Akselrod, S., Richards, J., Kenet, B., Kenet, R., Cohen, R. and Shannon, D. (1986), "Analysis of Heart Rate and Respiratory Patterns in Sudden Infant Death Syndrome Victims and Control Infants," *Pediatric Research*, Vol. 20, 1986, pp. 680–684.

Graseby Medical Ltd. (1988), "Respiration Monitor Type MR10,", Technical Service Manual No. SM 108, Colonial Way, Watford, Herts, WD2 4LG England, 1988.

Gray, R.M., Olshen, R.A., Ikeda, D., Cosman, P.C., Perlmutter, S., Nash, C. and Perlmutter, K. (1995), "Evaluating Quality and Utility in Digital Mammography," In *Proceedings of the IEEE International Conference on Image Processing*, IEEE Computer Society Press, Los Alamitos, CA, USA, 1995, pp. 5–8.

Gray, R.M., Olshen, R.A., Ikeda, D., Cosman, P.C., Perlmutter, S.M., Nash, C. and Perlmutter, K.O. (1996), "Measuring Quality in Computer-Processed Radiological Images," In *Conference Record of the Twenty-Ninth Asilomar Conference on Signals, Systems and Computers*, IEEE Computer Society Press, Los Alamitos, CA, USA, 1996, pp. 489–493.

Grossman, A. and Morlet, J. (1984), "Decompositions of Hardy functions into square integrable wavelets of constant shape," *SIAM J. Maths*, Vol. 15, 1984, pp. 723–736.

Grossmann, A., Kronland-Martinet, R. and Morlet, J. (1987), "Reading and Understanding Continuous Wavelet Transforms," In Jean-Michel Combes, A.G. and Tchamitchian, P. (Eds.), *Wavelets: Time-Frequency Methods and Phase Space*, Marseille, France, December 1987, pp. 2–20.

Guilleminault, C., Ariagno, R., Korobkin, R., Coons, S., Owen-Boeddiker, M. and Baldwin, R. (1981), "Sleep parameters and respiratory variables in 'near miss' sudden infant death syndrome infants," *Pediatrics*, Vol. 68, 1981, pp. 354–360.

Haidmayer, R., Kenner, T. and Kurz, R. (1980), "Paradoxical Ventilatory Response of Babies to Pure Oxygen (Author's Translation from German)," *Wiener Medizische Wocherschrift*, Vol. 130, No. 3, 1980, pp. 128–129.

Haidmayer, R., Kurz, R., Kenner, T., Wurm, H. and Pfeiffer, K. (1982a), "Physiological and Clinical Aspects of Respiration Control in Infants with Relation to the Sudden Infant Death Syndrome," *Klinische Wochenschrift*, Vol. 60, 1982, pp. 9–18.

Haidmayer, R., Pfeiffer, K., Kenner, T. and Kurz, R. (1982b), "Statistical Evaluation of Respiratory Control in Infants to Assess Possible Risk for the Sudden Infant Death Syndrome (SIDS)," *European Journal of Pediatrics*, Vol. 138, 1982, pp. 145–150.

Harper, R., Schechtman, V. and Kluge, K. (1987), "Machine Classification of Infant Sleep State Using Cardiorespiratory Measures," *Electroencephalography and Clinical Neurophysiology*, Vol. 67, No. 4, October 1987, pp. 379–387.

Haykin, S. (1983), *Communication Systems*, John Wiley & Sons, New York, NY, USA, 1983.

Healy, Jr., D.M. and Weaver, J.B. (1992), "Two Applications of Wavelet Transforms in Magnetic Resonance Imaging," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, March 1992, pp. 840–860.

Healy, Jr., D.M., Lu, J. and Weaver, J.B. (1995), "Two Applications of Wavelets and Related Techniques in Medical Imaging," *Annals of Biomedical Engineering*, Vol. 23, 1995, pp. 637–665.

Heeger, D.J. and Teo, P.C. (1995), "A Model of Perceptual Image Fidelity," In *Proceedings of the IEEE International Conference on Image Processing*, IEEE Computer Society Press, Los Alamitos, CA, USA, 1995, pp. 343–345.

Herrera, R., Sclabassi, R., Mingui, S. and Dahl, R. (1999), "Single Trial Visual Event–Related Potential EEG Analysis Using the Wavelet Transform," In *Proceedings of the First Joint BMES/EMBS Conference*, IEEE, Piscataway, NJ, USA, 1999, P. 947.

Hewertson, J., Poets, C., Samuels, M., Boyd, S., Neville, B. and Southall, D. (1944), "Epileptic Seizure-Induced Hypoxemia in Infants with Apparent Life-Threatening Events," *Pediatrics*, Vol. 94, 1944, pp. 148–156.

Hilton, M.L. and Ogden, R.T. (1997), "Data Analytic Wavelet Threshold Selection in 2-D Signal Denoising," *IEEE Transactions on Signal Processing*, Vol. 45, No. 2, February 1997, pp. 496–500.

Hodgman, J.E., Hoppenbrouwers, T., Geidel, S., Hadeed, A., Sterman, M.B., Harper, R. and McGinty, D. (1982), "Respiratory behaviour in near-miss sudden infant death syndrome," *Pediatrics*, Vol. 69, 1982, pp. 785–792.

Hunt, C.E. and Brouillette, R.T. (1987), "Sudden infant death syndrome: 1987 perspective," *Journal of Pediatrics*, Vol. 110, No. 5, 1987, pp. 669–677.

Hunt, C.E., Brouillette, R.T., Liu, K. and Klemka, L. (1985), "Day-to-day pneumogram variability," *Pediatric Research*, Vol. 19, 1985, pp. 174–177.

Jahnukainen, T., Lindqvist, A., Jalonen, J. and Vèalimèaki, I. (1995), "Skin Blood Flow Oscillations Respond More Effectively to Rhythmic Thermal Stimulation Than to Continuous or Periodic Breathing in Newborn Infants," *Early Human Development*, Vol. 42, No. 3, August 1995, pp. 195–207.

Kahn, A., Groswasser, J., Rebuffat, E., Sottiaux, M., Blum, D., Foerster, M., Franco, P., Bochner, A., Alexander, M., Bachy, A., Richard, P., Verghote, M., Polain, D.L. and Wayenberg, L. (1992), "Sleep and cardiorespiratory characteristics of infant victims of sudden death: a prospective case-control study," *Sleep*, Vol. 15, No. 4, 1992, pp. 287–292.

Kaiser, G. (1994), *A Friendly Guide to Wavelets*, Birkhäuser, 1994.

Kang, S.C., Lee, S.M. and Hong, S.H. (1998), "Noise Reduction of Echocardiographic Images Using Wavelet Filtering," In *Proceedings of ICSP'98*, IEEE, 1998, pp. 267–270.

Karlsson, G. and Vetterli, M. (1990), "Theory of Two-Dimensional Multirate Filter Banks," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 38, No. 6, June 1990, pp. 925–934.

Katz-Solomon, M. and Milerad, J. (1996), "Ventilatory and heart-rate responses to moderate CO_2-loading in infants at risk of SIDS," In *Conference Program of the Fourth SIDS International Conference*, SIDS International, Washington, USA, 23-26 June 1996, pp. 128–129.

Kelly, D.H., Walker, A.M., Cahen, L. and Shannon, D.C. (1980), "Periodic Breathing in Siblings of Sudden Infant Death Syndrome Victims," *Pediatrics*, Vol. 66, No. 4, October 1980, pp. 515–520.

Kelly, D.H., Golub, H., Carley, D. and Shannon, D.C. (1986), "Pneumograms in infants who subsequently died of sudden infant death syndrome," *Journal of Pediatrics*, Vol. 109, 1986, pp. 249–254.

Kempe, C.H., Silver, H.K. and O'Brien, D. (1974), *Current paediatric diagnosis and treatment*, Lange Medical Publications, Los Altos, California, 3rd ed., 1974.

Kenney, A.R. and Chapman, S. (1996), *Benchmarking Image Quality: From Conversion to Presentation*, Ithaca, NY: Department of Preservation and Conservation, Cornell University, 1996.

Keselbrener, L. and Akselrod, S. (1996), "Selective Discrete Fourier Transform Algorithm for Time-Frequency Analysis: Method and Application on Simulated and Cardiovascular Signals," *IEEE Transactions on Biomedical Engineering*, Vol. 43, No. 8, August 1996, pp. 789–802.

Kovačević, J. and Vetterli, M. (1992), "Nonseparable Multidimensional Perfect Reconstruction Filter Banks and Wavelet Bases for $\mathcal{R}^n$," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, March 1992, pp. 533–555.

Kovačević, J. and Vetterli, M. (1993), "FCO Sampling of Digital Video Using Perfect Reconstruction Filter Banks," *IEEE Transactions on Image Processing*, Vol. 2, No. 1, January 1993, pp. 118–122.

Kovačević, J. and Vetterli, M. (1995), "Nonseparable Two- and Three-Dimensional Wavelets," *IEEE Transactions on Signal Processing*, Vol. 43, No. 5, May 1995, pp. 1269–1273.

Laine, A., Fan, J. and Yang, W. (1995), "Wavelets for Contrast Enhancement of Digital Mammography," *IEEE Engineering in Medicine and Biology*, Vol. 14, No. 5, September/October 1995, pp. 536–550.

Leachtenauer, J.C., Malila, W., Irvine, J., Colburn, L. and Salvaggio, N. (1997), "General Image-Quality Equation: GIQE," *Applied Optics*, Vol. 36, No. 32, November 1997, pp. 8332–8328.

Lee, D.T. and Yamamoto, A. (1994), "Wavelet Analysis: Theory and Applications," *Hewlett-Packard Journal*, Vol. 45, No. 6, December 1994, pp. 44–54.

Lee, D., Caces, R., Kwiatkowski, K., Cates, D. and Rigatto, H. (1987), "A developmental study on types and frequency distribution of short apnoeas (3 to 15 seconds) in term and preterm infants," *Pediatric Research*, Vol. 22, No. 3, 1987, pp. 344–349.

Lewis, A. and Knowles, G. (1992), "Image Compression Using the 2-D Wavelet Transform," *IEEE Transactions on Image Processing*, Vol. 1, No. 2, April 1992, pp. 244–250.

Liu, J. and Moulin, P. (1997), "Complexity-Regularized Image Denoising," In *Proceedings International Conference on Image Processing*, IEEE Computer Society, Los Alamitos, CA, USA, 1997, pp. 370–373.

Liu, J. and Moulin, P. (1999), "Image Denoising Based on Scale-Space Mixture Modeling of Wavelet Coefficients," In *Proceedings 1999 International Conference on Image Processing*, IEEE, Piscataway, NJ, USA, October 1999, pp. 386–390.

Macey, P. (1998), *Apnoea Detection*, PhD thesis, University of Canterbury, Christchurch, New Zealand, January 1998.

Macey, K.E. and Page, W.H. (2000), "A Method for Selecting a Wavelet to Analyze Infant Breathing Signals," *Submitted to: Applied Signal Processing*, 2000.

Macey, P.M., Ford, R.P.K., Brown, P.J., Larkin, J., Fright, R.W. and Garden, K. (1995), "Apnoea detection: human performance and reliability of a computer algorithm," *Acta Paediatricia*, Vol. 84, 1995, pp. 1103–1107.

Macey, K., Page, W., Harper, R., Macey, P. and Ford, R. (2000a), "Calculating Rhythmicity of Infant Breathing Using Wavelets," In *Proceedings of SPIE Conference on Wavelet Applications in Signal and Image Processing VIII*, July 2000. Vol. 4119, December 2000.

Macey, K.E., Harper, R.M., Page, W.H. and Macey, P.M. (2000b), "Quantifying Amplitude Modulation of Breathing in Infants Using the Wavelet Transform," In *CDROM Proceedings of the World Congress on Medical Physics and Biomedical Engineering*, paper 5892-11144, 4 pages, July 2000.

Mallat, S. (1989), "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, July 1989, pp. 674–693.

Manjunath, B. and Ma, W. (1996), "Texture Features for Browsing and Retrieval of Image Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, August 1996, pp. 837–841.

McDicken, W. (1991), *Diagnostic Ultrasonics — Principles and use of Instruments*, Churchill Livingston, Medical Division of Longman Group UK Ltd, 3^{rd} ed., 1991.

Meyer, Y. (1993), *Wavelets: Algorithms and Applications*, SIAM, Pennsylvania, 1993.

Milne, A.D. and Ruggins, N. (1989), "Sudden infant death syndrome: recent focus on the respiratory system," *British Medical Journal*, Vol. 298, 1989, pp. 689–690.

Misiti, M., Misiti, Y., Oppenheim, G. and Poggi, J.M. (1996), *Wavelet Toolbox User's Guide*, The MathWorks, Inc., Natick, Massachusettes, USA, 1996.

Mojsilović, A., Marković, S. and Popović, M. (1997), "Characterisation of Visually Similar Diffuse Diseases From B-Scan Liver Images With the Nonseparable Wavelet Transform," In *Proceedings of the IEEE International Conference on Image Processing*, 1997, pp. 547–550.

Mojsilovic, A., Rackov, D. and Popovic, M. (1998), "On the Selection of an Optimal Wavelet Basis for Texture Characterization," In *Proceedings of 1998 International Conference on Image Processing*, IEEE Computer Society, Los Alamitos, CA, USA, October 1998, pp. 678–682.

Moulin, P. and Liu, J. (1999), "Analysis of Multiresolution Image Denoising Schemes Using Generalized Gaussian and Complexity Priors," *IEEE Transactions on Information Theory*, Vol. 45, No. 3, April 1999, pp. 909–919.

Muraki, S. (1993), "Volume Data and Wavelet Transforms," *IEEE Computer Graphics and Applications*, July 1993, pp. 50–56.

Nelson, E.A.S., Williams, S.M., Taylor, B.J., Morris, B. and Ford, R.P.K. (1989), "Postneonatal mortality in south New Zealand: necropsy data review," *Paediatric and Perinatal Epidemiology*, Vol. 3, 1989, pp. 375–385.

Ni, H., Zhang, J., Glotzbach, S., Schechtman, V. and Harper, R. (1994), "Dynamic Respiratory Responses to Preoptic/Anterior Hypothalamic Warming in the Sleeping Cat," *Sleep*, Vol. 17, No. 8, 1994, pp. 657–664.

Nicolier, F., Laligant, O., Truchetet, F., Legrand, A. and Kuhler, S. (1999), "Human Cell Texture Analysis with Quincunx Spline Wavelet Transform," In *Proceedings of the SPIE, Vol. 3644 (Human Vision and Electronic Imaging IV, San Jose, CA, USA, 25 - 28 Jan, 1999)*, SPIE - International Society of Optical Engineering, 1999, pp. 606–615.

Nill, N.B. and Bouzas, B.H. (1992), "Objective Image Quality Measure Derived from Digital Image Power Spectra," *Optical Engineering*, Vol. 31, No. 31, April 1992, pp. 813–825.

Odegard, J.E., Gopinath, R.A. and Burrus, C. (1992), "Optimal Wavelets for Signal Decomposition and the Existence of Scale-Limited Signals," In *IEEE Proceedings of the International Conference on Acoustics, Speech and Signal Processing - 92*, San Francisco, California, USA, 1992.

Oren, J., Kelly, D. and Shannon, D. (1986), "Identification of a high-risk group for Sudden Infant Death Syndrome among infants who were resuscitated for sleep apnea," *Pediatrics*, Vol. 77, No. 4, 1986, pp. 495–499.

Panigrahy, A., Filiano, J., Sleeper, L., Mandell, F., Valdes-Dapena, M., Kraus, H., Rave, L., Foley, E., White, W. and Kinney, H. (In Press), "Decreased Serotonergic Receptor Binding in Rhombic-Lipped Derived Regions of the Medulla Oblongata in the Sudden Infant Death Syndrome," *J Neuropathol Exp Neurol*, In Press.

Perlmutter, S.M., Cosman, P.C., Gray, R.M., Olshen, R.A., Ikeda, D., Adams, C.N., Betts, B., Williams, M., Perlmutter, K.O., Li, J., Aiyer, A., Fajardo, L., Birdwell, R. and Daniel, B.L. (1997), "Image Quality in Lossy Compressed Digital Mammograms," *IEEE Transactions on Signal Processing*, Vol. 59, No. 2, June 1997, pp. 189–210.

Preiss, G., Iscoe, S. and Polosa, C. (1975), "Analysis of a Periodic Breathing Pattern Associated with Mayer Waves.," *American Journal of Physiology*, Vol. 228, No. 3, March 1975, pp. 768–774.

Rantonen, T., Jalonen, J., Grönlund, J., Antila, K., Southall, D. and Välimäki, I. (1998), "Increased Amplitude Modulation of Continuous Respiration Precedes Sudden Infant Death Syndrome - Detection by Spectral Estimation of Respirogram," *Early Human Development*, Vol. 53, 1998, pp. 53–63.

Richard, C., Mosko, S. and McKenna, J. (1998), "Apnea and Periodic Breathing in Bed-Sharing and Solitary Sleeping in Infants," *Journal of Applied Physiology*, Vol. 84, No. 4, April 1998, pp. 1374–1380.

Richards, J., Alexander, J., Shinebourne, E., de Swiet, M., Wilson, A. and Southall, A. (1984), "Sequential 22-Hour Profiles of Breathing Patterns and Heart Rate in 110 Full-Term Infants During Their First 6 Months of Life," *Pediatrics*, Vol. 74, No. 5, November 1984, pp. 763–777.

Richardson, Jr., W.B. (1995), "Applying Wavelets to Mammograms," *IEEE Engineering in Medicine and Biology*, Vol. 14, No. 5, September/October 1995, pp. 551–560.

Rioul, O. (1993), "Regular Wavelets: A Discrete-Time Approach," *IEEE Transactions on Signal Processing*, Vol. 41, No. 12, December 1993, pp. 3572–3579.

Rioul, O. and Vetterli, M. (1991), "Wavelets and Signal Processing," *IEEE Signal Processing Magazine*, October 1991, pp. 14–38.

Roy, M., Kumar, V.R., Kulkarni, B., Sanderson, J., Rhodes, M. and Stappen, M.V. (1999), "Simple Denoising Algorithm Using Wavelet Transform," *AIChE J.*, Vol. 45, No. 11, 1999, pp. 2461–2466.

Ruhsam, C., Pfeutzner, H., Nopp, P., Nakesch, H. and Frais-Keolbl, H. (1995), "Time-Frequency Spectral Analysis of Electric Field Plethysmography Signals," *Medical Progress Through Technology*, Vol. 21, No. 1, 1995, pp. 17–28.

Russ, J.C. (1998), *The Image Processing Handbook*, CRC Press in cooperation with IEEE Press, Forida, USA, 3rd ed., 1998.

Schechtman, V. and Harper, R. (1991), "Time of Night Effects on Heart Rate Variation in Normal Neonates," *Journal of Developmental Physiology*, Vol. 16, No. 6, December 1991, pp. 349–353.

Schechtman, V., Harper, R., Kluge, K., Wilson, A., Hoffmann, H. and Southall, D. (1988), "Cardiac and Respiratory Patterns in Normal Infants and Victims of the Sudden Infant Death Syndrome," *Sleep*, Vol. 11, No. 5, 1988, pp. 413–424.

Schechtman, V., Harper, R. and Harper, R. (1994), "Distribution of Slow-Wave EEG Activity Across the Night in Developing Infants," *Sleep*, Vol. 17, 1994, pp. 316–322.

Schechtman, V., Lee, M., Wilson, A. and Harper, R. (1996), "Dynamics of Respiratory Patterning in Normal Infants and Infants Who Subsequently Died of the Sudden Infant Death Syndrome," *Sleep*, Vol. 40, No. 4, 1996, pp. 571–577.

Scheffer, F., Stute, H., Sontheimer, D., Meissner, A. and Linderkamp, O. (1996), "Are numbers of apneas, bradycardias, or tachycardias indicators for a higher risk of SIDS in preterm infants with BPD?," In *Conference Program of the Fourth SIDS International Conference*, SIDS International, Washington, USA, 1996, pp. 130–131.

Scholz, W. (1963), *Selective Vulnerability of the Brain in Hypoxaemia* eds. J.P Scadé and W.H. McMenemey, F.A. Davis Co, Phildelphia, PA, 1963, pp. 257–267.

Senhadji, L., Dillenseger, J.D., Wendling, F., Rocha, C. and Kinie, A. (1995), "Wavelet Analysis of EEG for Three Dimensional Mapping of Epiletic Events," *Annals of Biomedical Engineering*, Vol. 23, No. 5, September/October 1995, pp. 543–552.

Siegel, S. (1956), *Nonparametric Statistics For the Behavioral Sciences*, McGraw-Hill Book Company, Inc., New York, USA, 1956.

Smith, M.J.T. and Thomas P. Barnwell, I. (1987), "A New Filter Bank Theory for Time-Frequency Representation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 35, No. 3, March 1987, pp. 314–327.

Southall, D.P. (1988), "Role of apnoea in the sudden infant death syndrome: a personal view," *Pediatrics*, Vol. 80, 1988, pp. 73–84.

Stanton, A. (1984), "Sudden Infant Death. Overheating and Cot Death," *Lancet*, Vol. 2, No. 8413, 1984, pp. 1199–1201.

Stein, I. and Shannon, D.C. (1975), "The pediatric pneumogram: a new method of detecting and quantating apnea in infants," *Pediatrics*, Vol. 55, No. 5, 1975, pp. 599–603.

Stein, I.M., White, A., Kenndey, J.L., Merisalo, R.L., Chernoff, H. and Gould, J.B. (1979), "Apnea recodings of health infants at 40, 44, and 52 weeks postconception," *Pediatrics*, Vol. 63, No. 5, 1979, pp. 724–730.

Storm, H., Nylander, G. and Saugstad, O. (1999), "The Amount of Brainstem Gliosis in Sudden Infant Death Syndrome (SIDS) Victims Correlates with Maternal Cigarette Smoking During Pregnancy," *Acta Paediatricia*, Vol. 88, 1999, pp. 13–18.

Strang, G. and Nguyen, T. (1996), *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, Massachusetts, USA, 1996.

Strickland, R. and Hahn, H. (1997), "Wavelet Transform Methods for Object Detection and Recovery," *IEEE Transactions on Image Processing*, Vol. 6, No. 5, May 1997, pp. 724–735.

Sun, M. and Sclabassi, R. (1995), "Symmetric Wavelet Edge Detector of the Minimum Length," In *Proceedings of the IEEE International Conference on Image Processing*, IEEE Comput. Soc. Press, Los Alamitos, CA, USA, 1995, pp. 177–180.

Sun, M. and Sclabassi, R. (1998), "Precise Determination of Starting Time of Epileptic Seizures Using Subdural EEG and Wavelet Transforms," In *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, IEEE, New York, NY, USA, 1998, pp. 257–260.

Sun, M., Tsui, F.C. and Sclabassi, R.J. (1993), "Multiresolution Source Localization Using the Wavlet Transform," In Li, J.J. and Reisman, S. (Eds.),*Proceedings of the 1993 IEEE 19th Annual Northeast Bioengineering Conference*, IEEE, New York, NY, USA, 1993, pp. 88–91.

Surman, K., Carr, J., Columbi, Y., Fright, R. and Garden, K. (1993), "Development of a System for Three-Dimensional Graphics from Ultrasound," In *Proceedings of the First New Zealand Conference on Image and Vision Computing*, IRL, Auckland, New Zealand, 16–18 August 1993, pp. 367–373.

Tappin, D.M., Ford, R.P., Nelson, K.P., Price, B., Macey, P.M., Dove, R., Larkin, J. and Slade, B. (1996a), "Breathing, sleep state, and rectal temperature oscillations," *Archives of Disease in Childhood*, Vol. 74, 1996a, pp. 427–431.

Tappin, D.M., Ford, R.P., Nelson, K.P., Price, B., Macey, P.M. and Dove, R. (1996b), "Central apnoea is not increased in normal infants after vaccination," In *Conference Program of the Fourth SIDS International Conference*, SIDS International & SIDS Alliance, Washington, USA, 1996b, P. 119.

Tappin, D., Ford, R., Nelson, K., Price, B., Macey, P. and Dove, R. (1997), "The Febrile Stress of Routine Vaccination Does Not Increase Central Apnoea in Normal Infants," *Acta Paediatricia*, Vol. 86, 1997, pp. 873–880.

Taylor, C.C., Pizlo, Z., Allebach, J.P. and Bouman, C.A. (1997), "Image Quality Assessment with a Gabor Pyramid Model of the Human Visual System," In *Proceedings of the 1997 IS&T/SPIE International Symposium on Electronic Imaging and Technology*, SPIE, San Jose, CA, USA, February 1997, pp. 58–69.

Tsui, F.C., Sun, M., Li, C.C. and Sclabassi, R.J. (1995), "A Wavelet Based Neural Network for Prediction of ICP Signal," In *Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society*, IEEE, New York, NY, USA, 1995, pp. 1045–1046.

Tsui, F.C., Li, C.C., Sun, M. and Sclabassi, R.J. (1997), "A Comparative Study of Two Biorthogonal Wavelet Transforms in Time Series Prediction," In *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, IEEE, New York, NY, USA, 1997, pp. 1791–1796.

Tuffnell, C. (1993), *Biomedical Engineering Aspects of Infant Thermoregulation and Respiration*, PhD thesis, Electrical and Electronic Engineering Department, University of Canterbury, Christchurch. New Zealand, 1993.

Unser, M. and Aldroubi, A. (1996), "A Review of Wavelets in Biomedical Applications," *Proceedings of the IEEE*, Vol. 84, No. 4, April 1996, pp. 626–638.

Unser, M., Aldroubi, A. and Eden, M. (1992), "On The Asymptotic Convergence of B-Spline Wavelets to Gabor Functions," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, March 1992, pp. 864–872.

Vaidyanathan, P.P. (1990), "Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial," In *Proceedings of the IEEE*, January 1990, pp. 56–93.

Veldhuizen, T. (1998), *Grid Filters for Local Nonlinear Image Restoration*, Master's thesis, Indiana University Bloomington, 1998.

Vetterli, M. and Herley, C. (1992), "Wavelets and Filter Banks: Theory and Design," *IEEE Transactions on Signal Processing*, Vol. 40, No. 9, September 1992, pp. 2207–2232.

Vetterli, M. and Kovačević, J. (1995), *Wavelets and Subband Coding*, Prentice Hall, New Jersey, 1995.

Vetterli, M. and LeGall, D.J. (1989), "Perfect Reconstruction FIR Filter Banks: Some Properties and Factorizations," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 37, No. 7, July 1989, pp. 1057–1071.

Vetterli, M., Kovačević, J. and LeGall, D.J. (1990), "Perfect Reconstruction Filter Banks for HDTV Representation and Coding," *Signal Processing; Image communication 2*, 1990, pp. 349–363.

Villemoes, L.F. (1994), "Continuity of Nonseparable Quincunx Wavelets," *Applied and Computational Harmonic Analysis*, Vol. 1, 1994, pp. 180–187.

Wang, J.W., Chen, C.H., Chien, W.M. and Tsai, C.M. (1998), "Texture Classification Using Non-Separable Two-Dimensional Wavelets," *Pattern Recognition Letters*, Vol. 19, No. 13, November 1998, pp. 1225–1234.

Westen, S.J.P., Lagendijk, R.L. and Biemond, J. (1995), "Perceptual Image Quality Based on a Multiple Channel HVS Model," In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing - 95*, IEEE, New York, NY, USA, 1995, pp. 2351–2354.